# THE BELL SYSTEM
# Technical Journal

DEVOTED TO THE SCIENTIFIC AND ENGINEERING
ASPECTS OF ELECTRICAL COMMUNICATION

VOLUME XXXIX MARCH 1960 NUMBER 2

Fundamental Considerations in the Design of a Voice-Switched Speakerphone A. BUSALA 265

Integrated Magnetic Circuits for Synchronous Sequential Logic Machines U. F. GIANOLA 295

Amplitude Distribution of Shot Noise E. N. GILBERT AND H. O. POLLAK 333

On the Recovery of a Band-Limited Signal, After Instantaneous Companding and Subsequent Band Limiting H. J. LANDAU 351

Certain Mean Values in the Theory of the Traveling-Wave Amplifier L. A. MACCOLL 365

Radio Frequency Interference Considerations in the TD-2 Radio Relay System H. E. CURTIS 369

Diffused Junction Depletion Layer Calculations H. LAWRENCE AND R. M. WARNER, JR. 389

A Transversal Equalizer for Television Circuits R. V. SPERRY AND D. SURENIAN 405

Recent Bell System Monographs 423

Contributors to This Issue 428

COPYRIGHT 1960 AMERICAN TELEPHONE AND TELEGRAPH COMPANY

# THE BELL SYSTEM TECHNICAL JOURNAL

ADVISORY BOARD

H. I. ROMNES, *President, Western Electric Company*

J. B. FISK, *President, Bell Telephone Laboratories*

E. J. McNEELY, *Executive Vice President, American Telephone and Telegraph Company*

EDITORIAL COMMITTEE

A. C. DICKIESON, *Chairman*
S. E. BRILLHART
A. J. BUSCH
L. R. COOK
R. L. DIETZOLD
K. E. GOULD
E. I. GREEN
G. GRISWOLD, JR.
W. K. MACADAM
J. R. PIERCE
M. SPARKS
W. O. TURNER

EDITORIAL STAFF

W. D. BULLOCH, *Editor*
R. M. FOSTER, JR., *Assistant Editor*
C. POLOGE, *Production Editor*
J. T. MYSAK, *Technical Illustrations*
T. N. POPE, *Circulation Manager*

THE BELL SYSTEM TECHNICAL JOURNAL is published six times a year by the American Telephone and Telegraph Company, 195 Broadway, New York 7, N. Y. F. R. Kappel, President; S. Whitney Landon, Secretary; L. Chester May, Treasurer. Subscriptions are accepted at $5.00 per year. Single copies $1.25 each. Foreign postage is $1.08 per year or 18 cents per copy. Printed in U.S.A.

THE BELL SYSTEM
TECHNICAL JOURNAL

VOLUME XXXIX MARCH 1960 NUMBER 2

*Copyright 1960, American Telephone and Telegraph Company*

# Fundamental Considerations in the Design of a Voice-Switched Speakerphone

By A. BUSALA

(Manuscript received November 23, 1959)

*The speakerphone offers the advantage of hands-free telephony by replacing the familiar handset with a separate microphone and loudspeaker. The convenience of such an arrangement justifies extensive efforts to overcome its acoustic and transmission limitations. Voice switching overcomes some of the limitations but adds problems of its own. This study analyzes these problems and outlines a method that has proved helpful in the design of a voice-switched speakerphone and the evaluation of its performance.*

## I. INTRODUCTION

A speakerphone[1] is a telephone whose familiar handset, which places the transmitter close to the talker's lips and couples the receiver tightly to his ear, is replaced by a separate microphone and loudspeaker that can be set on a table a few feet from the user. This arrangement gives the customer definite advantages. In addition to leaving his hands free during a telephone conversation, it is a great help for physically handicapped persons and offers other convenient features, such as the possibility of having a small group join in the conversation and the avoidance of fatigue during lengthy calls.

However, these advantages are obtained at the price of some limitations.[2] To make up for the loss introduced by moving the instruments away from the head, gain is required in both the transmitting and the receiving paths. This gain is limited by a "singing" problem. A signal from the microphone reaches the loudspeaker via the sidetone path and comes back to the microphone through the acoustic coupling in the room. Too much gain in this loop causes singing. Even before reaching this condition, the loudspeaker-to-microphone acoustic coupling is the source of other undesirable effects. Incoming speech at the speakerphone end is fed back from the loudspeaker through the acoustic path to the microphone, whence it is returned to the distant talker with a certain delay. This is a form of talker echo. Similar reasoning would show that the distant party is also subjected to listener echo. Furthermore, when a speakerphone is used, the acoustic properties of the room and the ambient noise level are important, and often have adverse effects. This is in contrast to the performance of the regular telephone, whose instruments are so close to the user's head that transmission is largely independent of the surrounding conditions.

Voice switching is an answer to some of these limitations. Obviously, a voice-operated device that would apply to the speakerphone circuit the Vodas technique,[3] allowing only one direction of transmission to be fully active at a time, would eliminate both the singing problem and the talker and listener echoes. But the application of voice switching has problems and limitations of its own. There is the inherent limitation of one-way-at-a-time communication, which certainly detracts from the naturalness of the conversation. This is a penalty that one would have to pay, but is probably well worth the advantages gained. Then there are the problems connected with the switching operation. It is the purpose of this study to make a systematic analysis of all these problems and to determine their relation to the circuit design.

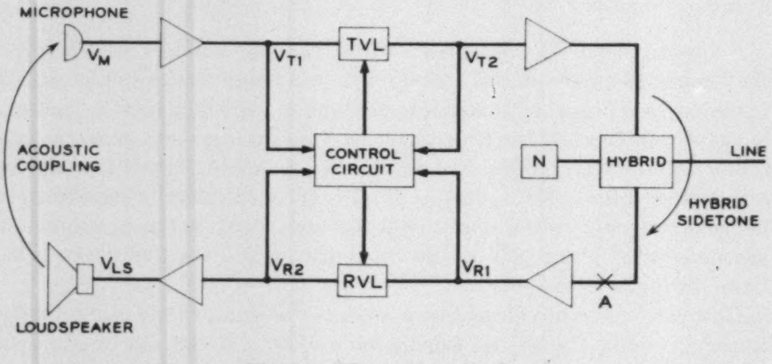

Fig. 1 — The master circuit.

## II. THE MASTER CIRCUIT

The general circuit of the essential elements of a voice-switched speakerphone is shown in Fig. 1, and will be referred to as the "master circuit". It consists of a microphone and a loudspeaker connected to a transmitting and a receiving branch with amplifiers to give the desired transmitting and receiving gains. A hybrid coil connects the transmitting and receiving branches to the line and to the sidetone-balancing network. With the practical limitations in sidetone balance and the inevitable acoustic coupling between loudspeaker and microphone, a substantial amount of loss must be inserted in the transmitting or the receiving branches to avoid singing. Therefore, a transmit variolosser, TVL, and a receive variolosser, RVL, have been included in Fig. 1. It is obvious that the loss will have to be switched* from one branch to the other, depending on whether the circuit is transmitting or receiving, and that one control circuit must operate simultaneously on both variolossers to insure that the minimum loss be always present.

The operation of the control circuit must be determined by some inputs. Four voltages whose ratio is not just a constant are available; these are two voltages, $V_{T1}$ and $V_{T2}$, in the transmitting branch and two voltages, $V_{R1}$ and $V_{R2}$, in the receiving branch. In general, all four voltages could be used simultaneously for control purposes.

## III. THE SUBMASTERS

The master circuit is general, but the use of four inputs to the control circuit makes it difficult to analyze its performance. Once the condition of the circuit (whether it is transmitting or receiving) and the amount of loss switched are known, the voltage $V_{T2}$ is related in a definite manner to $V_{T1}$, as is the voltage $V_{R2}$ to $V_{R1}$. This means that there are only two independent variables, and therefore two inputs — one voltage from the transmitting branch and one from the receiving branch — are sufficient for control purposes. There are four possible combinations of input voltages that can be used, and these correspond to four different circuit configurations, which will be called the "submasters". They are shown on Fig. 2.

If the control circuit is made intelligent enough, all the submasters can be made to give the same performance. Each one, however, has its own characteristics and will be examined individually later.

---

* While the term "switched gain" is often used in connection with a voice-operated device, the following description will be in terms of switched loss.

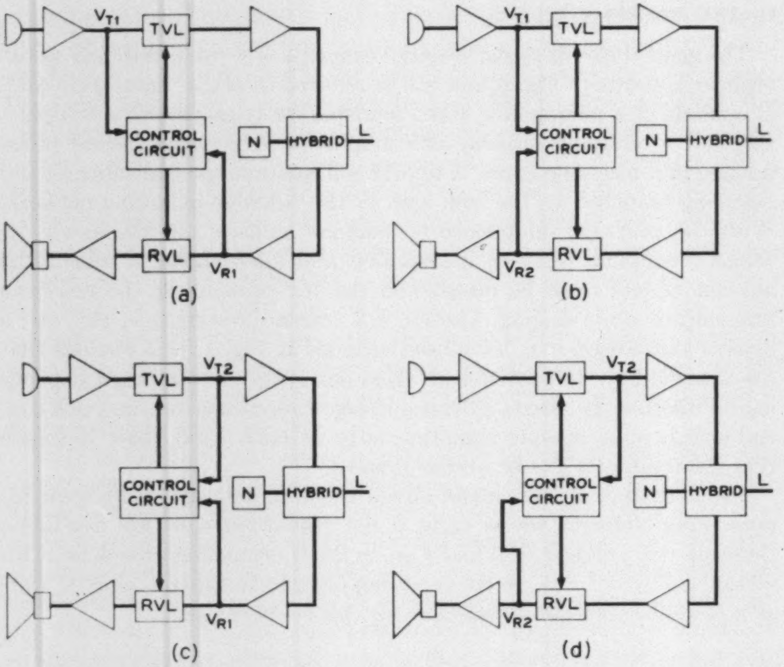
$V_{T1}$
TVL
CONTROL CIRCUIT
N
HYBRID
L
RVL
$V_{R1}$
(a)
$V_{T1}$
TVL
CONTROL CIRCUIT
N
HYBRID
L
RVL
$V_{R2}$
(b)
TVL
$V_{T2}$
CONTROL CIRCUIT
N
HYBRID
L
RVL
$V_{R1}$
(c)
TVL
$V_{T2}$
CONTROL CIRCUIT
N
HYBRID
L
RVL
$V_{R2}$
(d)

Fig. 2 — The submasters: (a) submaster #1; (b) submaster #2; (c) submaster #3; (d) submaster #4.

## IV. STEADY-STATE AND TRANSIENT PROBLEMS

The basic purpose of the voice-switching operation is to limit transmission (at least at full gain) to one direction at a time in order to avoid singing. It is the function of the control circuit to choose the direction of transmission on the basis of the two inputs and to operate accordingly on the variolossers. The critical parameters on which good performance of the speakerphone depends are the criterion applied by the control circuit to decide in which direction to have full gain and the time required to take the necessary action.

Starting, for instance, from a quiescent condition (which can be assumed to be the receiving one), the control circuit should switch the speakerphone into transmit as soon as there is a microphone input, in order not to lose the initial part of the signal. This implies that it would be desirable to have the control circuit very sensitive to the transmit input and to have a very short operating time. However, there are limitations to these requirements. For instance, if the circuit is too sensitive to microphone signals, room noise could easily impair reception of incoming signals. Also, if the control circuit were designed to switch back rapidly from transmit into receive every time the signal at the microphone fell below a certain threshold, the weaker sounds, especially consonants at the end of a word, would tend to be cut off and the transmitted speech would sound "choppy". Consequently, it seems desirable for the control circuit to possess a certain "hangover" in the transmitting condition after the signal has ceased. However, there are restrictions to the duration of this hangover. The longer the circuit remains transmitting, the more difficult it becomes for the distant party to break in. In other words, the initial part of an incoming signal might be lost. Therefore, the possibility of avoiding choppiness of the transmitted speech without impairing the received signal depends on the existence of satisfactory compromises for the sensitivity and the time constants of the control circuit.

From this typical example of switching operation, it is apparent that a good performance of the speakerphone is often a compromise in a number of characteristics, such as how fast it switches, how easily it stays in one condition once it has switched, how sensitive it is to room or line noise, or how easy it is to break in during a back-and-forth conversation.

In order to achieve the best compromise, it is helpful to classify the fundamental problems of voice switching into two groups, the steady-state and the transient problems. The ones in the first group concern the performance of the circuit in each condition, transmitting and receiving, with steady or slowly varying input signals. The ones in the second group concern its performance under transient conditions during the short time intervals when the input signals are rapidly varying.

The steady-state problems are:

i. Singing. This occurs, as already mentioned, when not enough loss is present in the transmitting or in the receiving branches.

ii. Transmit Blocking. In this case, the circuit is in the receive condition when it should be in the transmit condition. It can result from one of two conditions: a steady noise coming from the line may prevent the circuit from going into transmit, even with large signals into the microphone (transmit blocking due to noise operation); or a poor sidetone balance in the hybrid may generate a sidetone voltage large enough to hold the circuit in the receive condition (transmit blocking due to false switching).

iii. Receive Blocking. This problem is analogous to transmit blocking, and can also result from one of two conditions: a steady room noise may prevent the circuit from going into receive (receive blocking due to noise operation); or the acoustic coupling between loudspeaker and microphone may generate a pressure at the microphone large enough to hold the circuit in the transmit condition (receive blocking due to false switching).

The transient problems can be classified as follows:

i. Initial Clipping. This is the loss of the first part of a speech signal, and can occur either in transmitting or in receiving. It is, in general, a function of the operating time of the control circuit and switch.

ii. Final Clipping. This is the loss of the last part of a speech signal, and also can happen either in transmitting or in receiving. It is a function of the hangover time of the control circuit and switch.

iii. Echo. This problem is due to the finite decay time of the sound in a reverberant room. The sound put out by the loudspeaker is sustained in the room for a certain time after the incoming signal has ceased. This sound may be picked up and retransmitted by the microphone, giving the effect of a short, transient echo to the far end. The symmetrical problem of an echo effect in transmitting rather than in receiving is much subordinated, since there is no delay in the sidetone path through the hybrid, and any echo coming back from the line is very likely to be below a troublesome level.

## V. THE SWITCHING DIAGRAM

To analyze the performance of a speakerphone, it is helpful to use a plot that will be called the *switching diagram*. This graphical method offers the advantage over an algebraic analysis of showing at a glance the condition of the speakerphone and the margins against singing and blocking for any combination of input voltages. Furthermore, the diagram can be used to represent the effects of nonlinearity in the different parts of the circuit and the performance of the speakerphone during transients. In one convenient form of the diagram, the abscissa represents the voltage at the loudspeaker terminals and the ordinate the voltage across the microphone, with both voltages being on a logarithmic scale.

Rather than studying a single frequency, let us consider speech, assuming that the plotted voltages are weighted averages of some kind representing the speech power in the frequency band of interest. For simplicity, it will also be assumed that all the elements of the circuit have a flat response. Frequency shaping can be used advantageously, especially in eliminating some particular problems, such as receive blocking due to false switching and echo. Its effect could be analyzed in a similar manner by extending the application of the switching diagram.

### 5.1 *The Acoustic Coupling and the Hybrid Sidetone*

In the switching diagram one can plot two lines, which will be called the *acoustic coupling* and the *hybrid sidetone* [Fig. 3(a)]. The first represents the voltage that appears at the microphone terminals for a given voltage at the loudspeaker because of the acoustic coupling between loudspeaker and microphone. If all elements are linear, this is a 45° line whose position is a function of the acoustic environment, the distance between microphone and loudspeaker and the efficiencies of the two instruments. This line will be plotted for the limiting conditions in which we expect the speakerphone to operate, i.e., minimum distance between microphone and speaker and minimum room constant.[4] If one allows for overloading

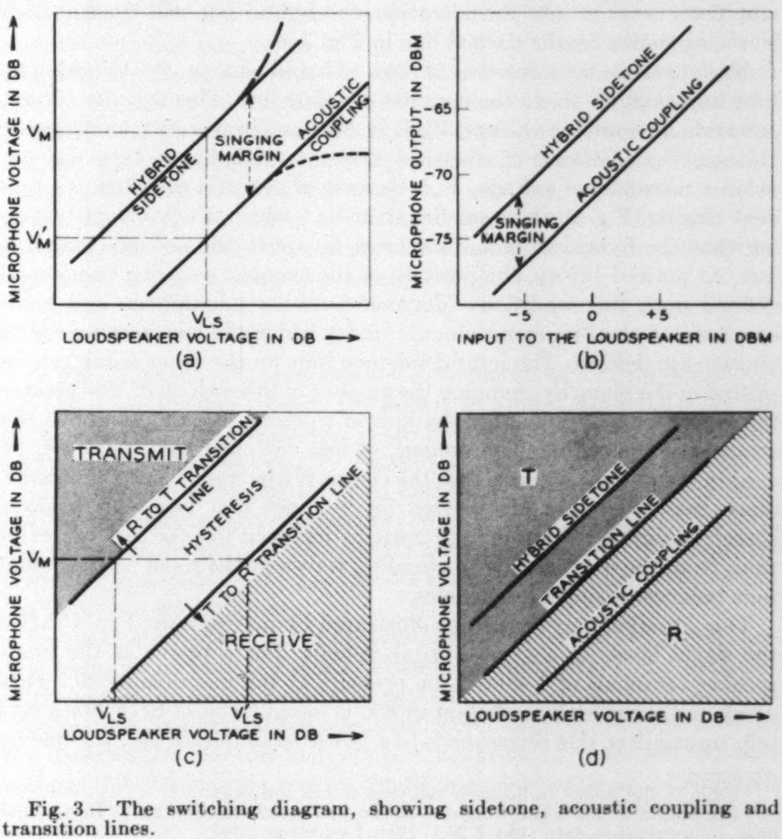
MICROPHONE VOLTAGE IN DB
$V_M$
$V'_M$
HYBRID SIDETONE
SINGING MARGIN
ACOUSTIC COUPLING
$V_{LS}$
LOUDSPEAKER VOLTAGE IN DB
(a)
MICROPHONE OUTPUT IN DBM
-65
-70
-75
HYBRID SIDETONE
ACOUSTIC COUPLING
SINGING MARGIN
-5
0
+5
INPUT TO THE LOUDSPEAKER IN DBM
(b)
MICROPHONE VOLTAGE IN DB
TRANSMIT
R TO T TRANSITION LINE
HYSTERESIS
T TO R TRANSITION LINE
RECEIVE
$V_M$
$V_{LS}$
$V'_{LS}$
LOUDSPEAKER VOLTAGE IN DB
(c)
MICROPHONE VOLTAGE IN DB
T
HYBRID SIDETONE
TRANSITION LINE
ACOUSTIC COUPLING
R
LOUDSPEAKER VOLTAGE IN DB
(d)

Fig. 3 — The switching diagram, showing sidetone, acoustic coupling and transition lines.

of the instruments, the acoustic coupling line will, in general, bend at high levels, as shown by the dotted line in Fig. 3(a).

The other line, the hybrid sidetone, represents the voltage appearing at the loudspeaker for a given voltage at the microphone. If all elements are linear, this is also a 45° line whose position is determined by the transmitting and receiving gain, the sidetone balance and the amount of loss switched. In general, the amount of loss switched in the transmitting branch may be different from the loss switched in the receiving branch. In this case, we would have two sets of hybrid sidetone lines, one obtained with the speakerphone in the transmit condition and one with it in the receive condition. The hybrid sidetone line for the poorest sidetone balance to be expected is plotted. If overloading of transmit and receive amplifiers is taken into consideration, the hybrid line will bend at high levels, as shown by the dashed line in Fig. 3(a).

An immediate consideration is that, to avoid singing, the hybrid sidetone line must lie above the acoustic coupling line. This is easily shown. A certain microphone voltage, $V_M$ [Fig. 3(a)], will generate a loudspeaker voltage, $V_{LS}$, because of electrical sidetone, and this in turn will develop a microphone voltage, $V_M'$, because of acoustic coupling. To prevent singing, $V_M'$ must be smaller than $V_M$, which is equivalent to saying that the hybrid sidetone line must lie above the acoustic coupling line. As noticed before, the position of the acoustic coupling line is well defined once the conditions (distance between microphone and loudspeaker and acoustic environment) under which the speakerphone is to operate are defined. The hybrid sidetone line, on the other hand, can be shifted in the plane by changing the amount of loss switched. The greater the loss switched, the farther the hybrid sidetone line moves above the acoustic coupling line. The amount of loss switched can, therefore, be used in the design to insure that the circuit is free from singing problems. Since these two lines are plotted for voltages that represent average speech energy, a certain margin must be included to take care of deviations from the average due to standing waves in the room or poor sidetone balance at single frequencies.

This situation may be further illustrated by an example. Fig. 4 shows the signal levels in a typical speakerphone. For a power at the loudspeaker terminals of 1 dbm,* the pressure at the microphone 30 inches away in a room with a constant of 800 is found to be 71 SL.† For a certain transmitter this corresponds to a power available at the microphone

---

* Dbm is the power in decibels referred to one milliwatt.

† SL, sound level, is the weighted sound pressure level at a point in a sound field, in accordance with ASA Z 24.1, 1951, Definition 1.380.

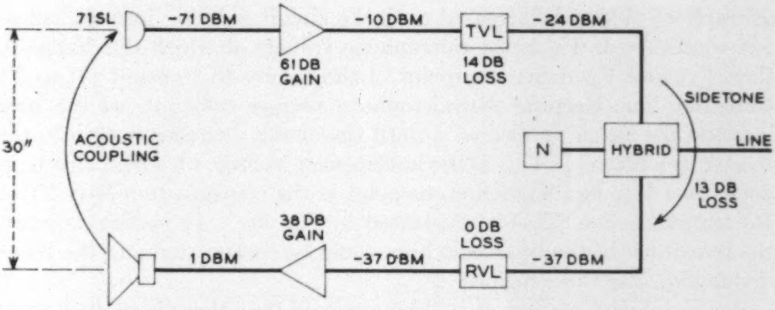

**Fig. 4 — Signal levels in a typical speakerphone, illustrating the singing problem**

terminals of −71 dbm. These power values define the position of the acoustic coupling line shown in Fig. 3(b). With a gain in the transmitting branch of 61 db, a gain in the receiving branch of 38 db and a hybrid loss in the coil of 13 db, it can be seen that the minimum loss to be switched is 14 db. In Fig. 4 the circuit is shown in the receive condition and the minimum loss of 14 db is inserted in the transmit variolosser. If the speakerphone switched this amount of loss, the hybrid sidetone line would lie exactly on top of the acoustic coupling line. If 20 db were switched, there would be a margin of 6 db, and the two lines would be separated by this amount, as shown in Fig. 3(b).

### 5.2 *The Transition Lines*

In Fig. 3(a) each point represents a combination of a microphone voltage and a loudspeaker voltage. For each such combination, the condition of the speakerphone, depending on the action of the control circuit, can be uniquely defined as being in transmit or receive, or it can be indeterminate, depending on the previous history of the particular combination of microphone and loudspeaker voltages. As shown in Fig. 3(c), the plane can be divided accordingly into three regions, transmit, receive and hysteresis, by two transition lines, the transmit-to-receive and the receive-to-transmit lines.

To make clear the meaning of these three regions and the two transition lines, let us take the master circuit shown in Fig. 1 and open the sidetone and the acoustic coupling paths by opening the circuit at point A in the receiving branch and acoustically isolating the microphone from the loudspeaker. We can then apply a signal at A toward the left that will put the circuit in the receive condition and generate an independent voltage $V_{LS}$ at the loudspeaker. With this voltage kept constant, the microphone voltage is increased until the circuit switches into the transmit condition. If $V_M$ is the microphone voltage at which this happens, then $V_{LS}$ and $V_M$ define one point of the receive-to-transmit (R to T) transition line. Keeping the microphone voltage constant, we can now increase the signal applied at A until the circuit switches back into the receive condition. If $V_{LS'}$ is the loudspeaker voltage at which this happens, then $V_{LS'}$ and $V_M$ define one point of the transmit-to-receive (T to R) transition line. This is illustrated in Fig. 3(c). In similar manner, the remainder of the transition lines could be constructed and the plane divided into the three regions.

For simplicity, at first, it will be assumed that there is no hysteresis region. This means that, for each point of the diagram, the control circuit uniquely defines the condition of the speakerphone. In this case, the two transition lines coincide, and the plane is divided into just two regions, transmit and receive.

The position and shape of the transition line in the diagram is completely determined by the characteristics of the circuit that controls the voice swtiching, and can be chosen by proper design. It is easy to show that, in order to avoid blocking problems, the transition line must lie in the diagram between the hybrid sidetone and the acoustic coupling lines, so that the hybrid sidetone line will be in the transmit region and the acoustic coupling line will be in the receive region, as shown in Fig. 3(d). Let us assume that this condition is not satisfied and that at least a portion of the hybrid sidetone line lies in the receive region of the diagram. This means that, for some values of the microphone voltage, the electrical sidetone through the hybrid develops a voltage at the loudspeaker large enough to put the speakerphone in the receive condition. This is the problem previously called transmit blocking due to false switching. Similarly, if the acoustic coupling line lies in the transmit region, the condition of receive blocking due to false switching develops.

As an example, in Fig. 5(a) the transition line is shown crossing the hybrid sidetone line at point A. For any microphone voltage greater than $V_M$, the loudspeaker voltage due to the sidetone path through the hybrid is large enough to put the circuit in the receive condition. The transmit region is limited at the upper end by the horizontal line through A. If the user of the speakerphone talks at a level greater than the level corresponding to the voltage $V_M$, he is "blocked out"; that is, the circuit stays in the receiving condition and his talking louder, which is the natural reaction, will not place the circuit into transmit.

A similar situation is shown in Fig. 5(b), where the transition line crosses the acoustic coupling line at B and the receive region is limited to

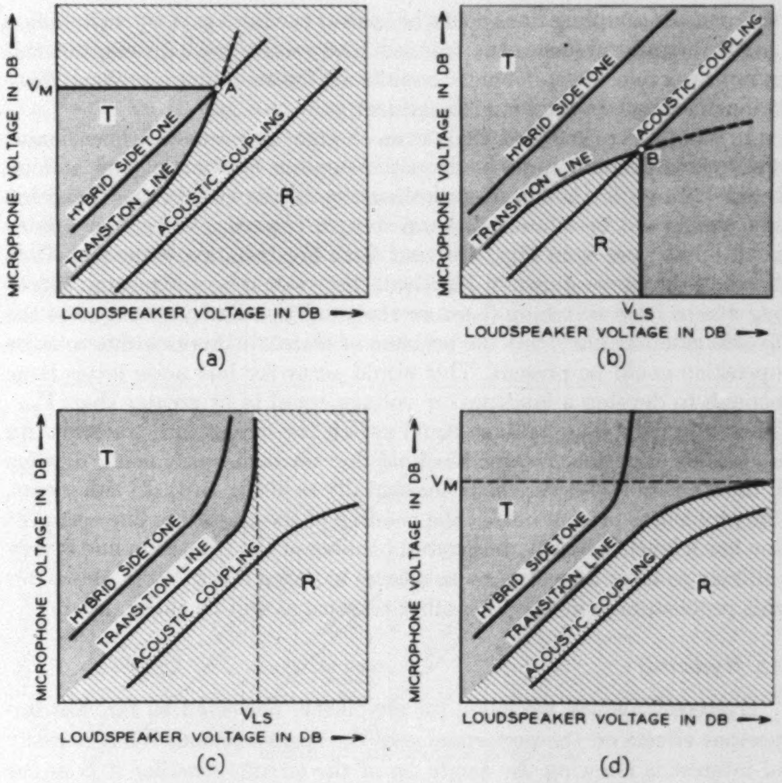
MICROPHONE VOLTAGE IN DB
LOUDSPEAKER VOLTAGE IN DB
$V_M$
A
T
R
HYBRID SIDETONE
TRANSITION LINE
ACOUSTIC COUPLING
(a)
B
$V_{LS}$
(b)
(c)
(d)

Fig. 5 — Switching diagrams showing blocking problems.

values of loudspeaker voltage smaller than $V_{LS}$. For values of loudspeaker voltage greater than $V_{LS}$, the circuit tends to go into transmit because of the acoustic coupling between loudspeaker and microphone. This condition has been referred to as receive blocking.

The requirement that the transition line lie between the sidetone and the acoustic coupling lines to avoid blocking problems is practically equivalent to the condition that the switching action be of a linear differential type; i.e., the greater the voltage present at the loudspeaker, the greater, by about the same amount, the microphone voltage necessary to switch the circuit into transmit. This is apparent in Fig. 3(d), where the transition line must be close to a 45° line in order to lie, in the range of interest, between the two other 45° lines. The sidetone and the acoustic coupling lines could be spread further apart by switching a greater amount of loss in the transmit and receive variolossers, but this is not desirable, since it tends to make all the switching problems more critical and voice switching less natural.

It should also be noted that, even though, as previously mentioned, the hybrid sidetone and the acoustic coupling lines may bend at high levels because of the effect of nonlinearity in the circuit, the transition line should not be allowed to bend sharply toward a vertical or a horizontal direction, as in Figs. 5(c) and 5(d). For instance, if the transition line had the shape shown in Fig. 5(c), there would be no transmit blocking due to false switching (because the transition line does not cross the hybrid sidetone line), but the problem of transmit blocking due to noise operation could be present. This would occur for line noise levels large enough to develop a loudspeaker voltage equal to or greater than $V_{LS}$. Then no microphone voltage could switch the circuit into transmit. An analogous situation, receive blocking due to room noise, could develop if the transition line were bent horizontally as in Fig. 5(d). In either case, the problem is present only if the bending of the transition line occurs at too low a level. Actually, horizontal bending of the transition line is very unlikely to be at a level to cause receive blocking due to room noise. But this problem may develop for other reasons, as will be shown later.

### 5.3 *Hysteresis*

Hysteresis, which we have, for simplicity, neglected so far, has important effects on the performance of the speakerphone. What is really of interest is knowing the condition of the circuit: whether it is in the transmitting, receiving or hysteresis region for each combination of the two independent variables, microphone voltage and incoming line voltage.

For this reason, it is useful at this point to make a change of coordinate in the switching diagram and use the incoming line voltage rather than the loudspeaker voltage on the abscissa. It should be pointed out that by incoming line voltage we mean the voltage due to an incoming signal and not merely the voltage appearing across the line terminals, since this could be the voltage due to an outgoing signal from the microphone output or a combination of an incoming signal and an outgoing signal. To make this change in coordinate, we translate the voltage at the loudspeaker into a line voltage by applying to it a correction that takes into consideration, besides the change in impedance, the gain in the receiving branch and the losses in the receive variolosser and the hybrid coil. In this way, a sidetone voltage is plotted in terms of an equivalent incoming line voltage that would give the same loudspeaker voltage.

With these assumptions, the sidetone and the acoustic coupling lines can be drawn on the new diagram. Starting with the acoustic coupling line, it now becomes necessary to define the condition of the circuit — whether it is in transmit or in receive — since the receive variolosser is in the path between the incoming line voltage and the voltage developed at the microphone. As shown in Fig. 6(a), the acoustic coupling line shifts by the amount of loss switched in the receive variolosser upon going from the receive to the transmit condition. Since linearity is assumed, the lines are straight at a 45° angle, and the shift can be considered to be either vertical or horizontal. When nonlinearities are present, as exempli-

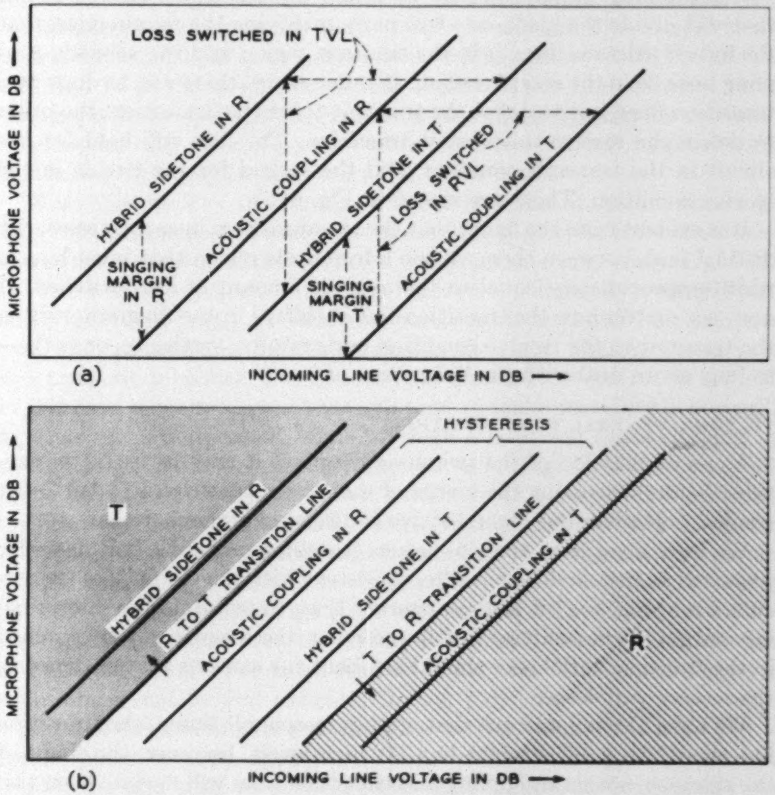
MICROPHONE VOLTAGE IN DB
LOSS SWITCHED IN TVL
HYBRID SIDETONE IN R
ACOUSTIC COUPLING IN R
HYBRID SIDETONE IN T
LOSS SWITCHED IN RVL
ACOUSTIC COUPLING IN T
SINGING MARGIN IN R
SINGING MARGIN IN T
(a)
INCOMING LINE VOLTAGE IN DB
MICROPHONE VOLTAGE IN DB
HYSTERESIS
T
HYBRID SIDETONE IN R
R TO T TRANSITION LINE
ACOUSTIC COUPLING IN R
HYBRID SIDETONE IN T
T TO R TRANSITION LINE
ACOUSTIC COUPLING IN T
R
(b)
INCOMING LINE VOLTAGE IN DB

Fig. 6 — Switching diagram with hysteresis.

fied by the dotted line in Fig. 3(a), the shift should be considered vertical if they occur in the circuit before the receive variolosser, and horizontal if they occur after the receive variolosser. Similarly, the hybrid sidetone line has two positions, which are shifted by the amount of loss switched in the transmit variolosser, with similar rules determining whether the shift should be considered horizontal or vertical.

In view of this, the two sets of lines in Fig. 6(a) correspond to the transmit and receive conditions. The line for hybrid sidetone in the transmit condition has been drawn below the line for acoustic coupling in the receive condition. This means that, when no loss is present in either the transmit or the receive variolosser, the circuit is not stable and will sing. If this were not true, there would be no reason to introduce voice switching. It follows that no single transition line can be drawn that will divide the plane into two parts satisfying the requirement that the hybrid sidetone lines lie in the transmit region and the acoustic coupling lines lie in the receive region. It is necessary, therefore, to draw two transition lines, one to define the transmit-to-receive transition, the other to define the receive-to-transmit transition. The first will hold for the circuit in the transmit condition and the second for the circuit in the receive condition. These are shown in Fig. 6(b).

It is evident from the figure that the two transition lines are necessarily distinct and, between them, define a hysteresis region that must have a width approximately equal to the average amount of loss switched. It does not matter how the transition line is shifted in the diagram — from the transmit to the receive condition horizontally, vertically, or both — as long as we deal with the linear part of it.

## VI. SOME GENERAL CONSIDERATIONS

As an application of the switching diagram, it may be useful to consider, before discussing the merits of each submaster circuit, what other general objectives one should strive for in speakerphone design.

We have mentioned that, to reduce switching problems, it is desirable to switch as little loss as possible, consistent with the stability of the circuit. From Fig. 6(a) it is also apparent that, to maintain the same singing margin in transmitting and in receiving, the amount of loss switched in the transmit variolosser should be about the same as that switched in the receive variolosser.

We have also pointed out that, within reasonable limits, the transition line should be a 45° straight line. At high levels, however, the shape of the sidetone, acoustic coupling and transition lines will deviate from this straight line. As long as the levels at which this happens are high compared with those of normal speech, the shape of these curves becomes immaterial. At low levels, on the other hand, it may be advantageous to have the transition lines deviate from the linear characteristic in a predetermined fashion. It has been recognized[5] that, when a voice-operated device selecting one of two directions of transmission on the basis of signal amplitude is located at a point where the signal-to-noise ratio coming from one direction is poorer than that coming from the opposite direction, it is advantageous to use an arrangement whereby the direction having the better signal-to-noise ratio is normally blocked and the direction having the poorer ratio is normally activated. In the case of a speakerphone, the incoming signal-to-noise ratio cannot be controlled and, therefore, can be poorer than the one from the near end. Thus, it is reasonable to design the control circuit so that, at low outgoing levels, the speakerphone is normally in the receive condition. This is equivalent to setting a threshold in the microphone voltage below which the circuit will only be in the receive condition. The transition line will then have the shape shown in Fig. 7(a). It is true that this shape implies transmit blocking for microphone voltages below the threshold, but, as previously mentioned, this may be desirable at very low levels.

When the shape of the transition line is like that of the one shown in Fig. 7(a), the manner in which the line is shifted in going from the transmit to the receive condition and vice versa becomes important. Fig. 7(b) shows a completely horizontal shift, Fig. 7(c) a completely vertical shift and Fig. 7(d) a partly vertical and partly horizontal shift. In the first case, the transmitting threshold remains the same whether the circuit is in transmit or receive. The choice of the threshold level is rather critical in this case, since too low a level will tend to make receiving "choppy", especially in a noisy room, while too high a level will cause clipping in transmitting. With a completely vertical shift [Fig. 7(c)] the threshold drops considerably when the circuit switches into transmit. It is possible, therefore, to choose the threshold level so that it is fairly low in the transmit condition and high in the receive condition, resulting in a very stable circuit, both in transmit and in receive. In this case, however, it might be difficult to switch from transmit into receive, since even a relatively low room noise might keep the circuit in transmit (receive blocking due to noise operation). For these reasons it may be desirable to choose a compromise such as that indicated in Fig. 7(d).

Another feature worth consideration is the volume control. For a speakerphone to work satisfactorily under widely varying room and line conditions, and also to suit different users' needs, it is necessary to have a volume control in the receiving branch. Adjusting the gain of the

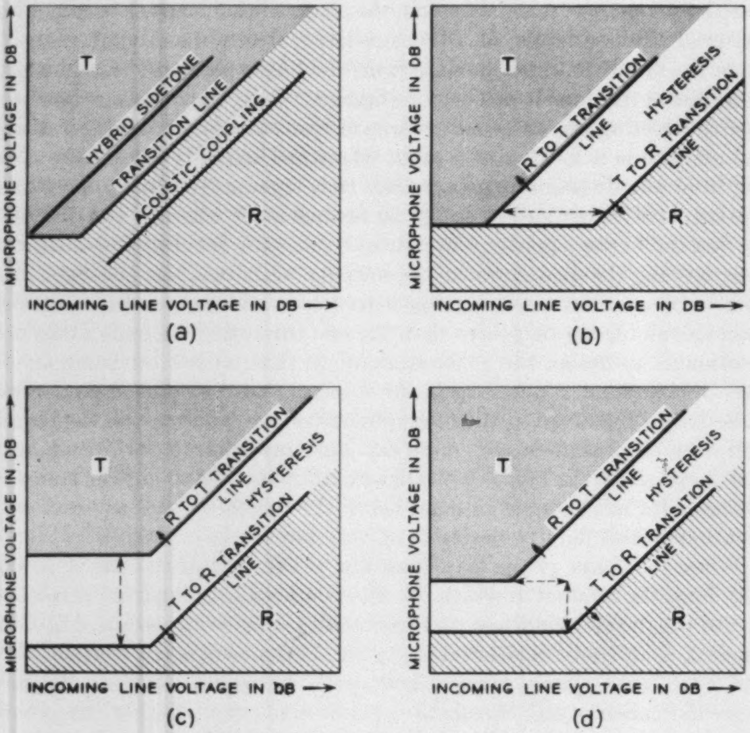
MICROPHONE VOLTAGE IN DB →
INCOMING LINE VOLTAGE IN DB →
T
R
HYBRID SIDETONE
TRANSITION LINE
ACOUSTIC COUPLING
R TO T TRANSITION LINE
T TO R TRANSITION LINE
HYSTERESIS
(a)
(b)
(c)
(d)

Fig. 7 — Switching diagram, transition lines with threshold.

receiving amplifier or inserting a loss in the receiving branch has the effect of shifting the acoustic coupling lines in the switching diagram without moving the hybrid sidetone lines. Since the circuit must be stable even with maximum gain setting, it follows that, as the volume control is turned down, the acoustic coupling lines move away from the hybrid sidetone lines, as shown in Fig. 8(a), thus unnecessarily increasing the amount of singing margins. Keeping in mind the objective of switching only the amount of loss necessary to make the circuit stable, it is desirable to arrange the volume control in such a way that, when turned down, it will reduce by the same amount the receiving level and the loss switched, in both the transmit and the receive variolossers.[6] This arrangement has the effect [shown in Fig. 8(b)] of shifting the sidetone and acoustic coupling lines for the receive condition toward the corresponding lines for the transmit condition. The lines for the transmit condition

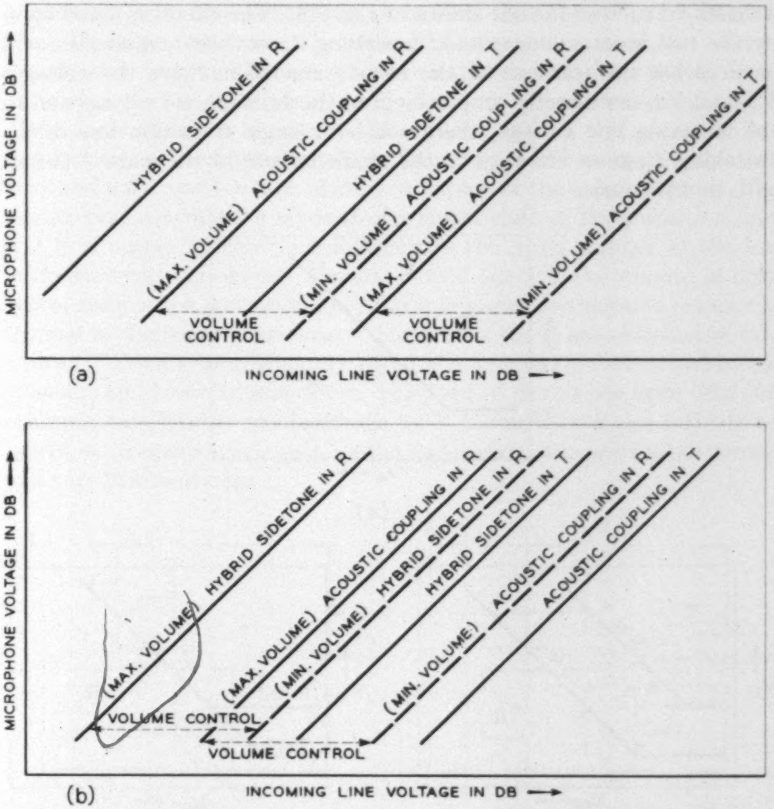
MICROPHONE VOLTAGE IN DB ⟶
HYBRID SIDETONE IN R
(MAX. VOLUME) ACOUSTIC COUPLING IN R
HYBRID SIDETONE IN T
(MIN. VOLUME) ACOUSTIC COUPLING IN R
(MAX. VOLUME) ACOUSTIC COUPLING IN T
(MIN. VOLUME) ACOUSTIC COUPLING IN T
VOLUME CONTROL
VOLUME CONTROL
(a)
INCOMING LINE VOLTAGE IN DB ⟶
MICROPHONE VOLTAGE IN DB ⟶
(MAX. VOLUME) HYBRID SIDETONE IN R
(MAX. VOLUME) ACOUSTIC COUPLING IN R
(MIN. VOLUME) HYBRID SIDETONE IN R
HYBRID SIDETONE IN T
(MIN. VOLUME) ACOUSTIC COUPLING IN R
ACOUSTIC COUPLING IN T
VOLUME CONTROL
VOLUME CONTROL
(b)
INCOMING LINE VOLTAGE IN DB ⟶

Fig. 8 — Switching diagram, two types of volume control: (a) keeping the amount of switched loss constant; (b) varying the amount of switched loss.

remain stationary and the margins stay constant. The control circuit, of course, must be so designed that, as the volume control is changed, the transition lines will always maintain the proper position with respect to the sidetone and acoustic coupling lines in the diagram.

## VII. FOUR SUBMASTER CIRCUITS

### 7.1 *Submaster #1*

A block diagram for this type of speakerphone is shown in Fig. 2(a), with the inputs to its control circuit being $V_{T1}$ and $V_{R1}$. Let us assume that the control circuit by itself has a simple linear characteristic with a threshold such as the one shown in Fig. 9(a). The circuit is acted upon by the two input voltages and, depending upon their magnitudes, will be in either the transmit or the receive condition. Since the voltages $V_{T1}$ and $V_{R1}$ are directly proportional to the microphone voltage and to the incoming line voltage, there will be a single transition line in the switching diagram identical to the characteristic of the control circuit with no hysteresis.

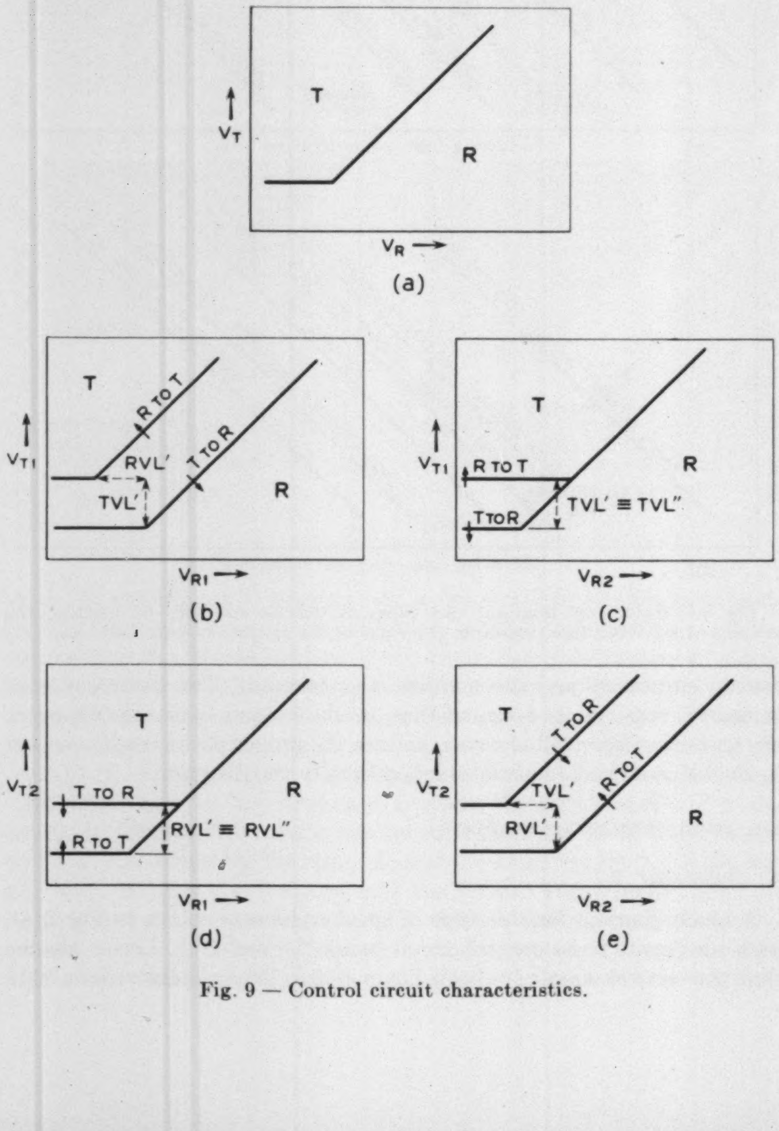
T
V_T
R
V_R
(a)
T
R TO T
V_T1
RVL′
T TO R
TVL′
R
V_R1
(b)
T
V_T1
R TO T
R
TVL′ ≡ TVL″
T TO R
V_R2
(c)
T
V_T2
T TO R
R
R TO T
RVL′ ≡ RVL″
V_R1
(d)
T
T TO R
V_T2
TVL′
R TO T
RVL′
R
V_R2
(e)

Fig. 9 — Control circuit characteristics.

From the previous discussions, it is clear that the lack of hysteresis gives blocking problems. Therefore, the characteristic of the control circuit must be somewhat more sophisticated and display some hysteresis, such as in Fig. 9(b). One way of obtaining this characteristic and of easily controlling the horizontal and vertical shift of the transition line is shown schematically in Fig. 10(a). Two extra variolossers are employed, TVL′ and RVL′, one for each of the two inputs of the control circuit. The loss of TVL′ controls the amount of vertical shift of the transition line, and it is varied by the control circuit in the same manner as the loss in the transmit variolosser. The loss in RVL′ controls the amount of horizontal shift of the transition line, and it is varied in the same manner as the loss in the receive variolosser. The sum of the losses switched in TVL′ and RVL′ should approximately equal the average loss switched in the transmit and receive variolossers. This type of circuit has been used, for example, in a British speakerphone by Winston Electronics Ltd.[7] It requires extra components such as the two variolossers shown but results in a very flexible design.

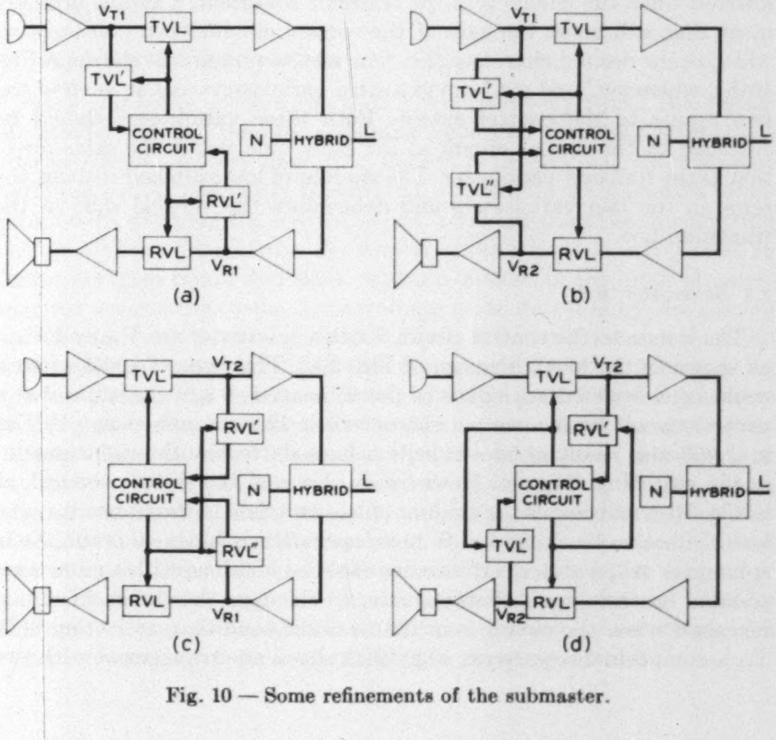
$V_{T1}$
TVL
TVL′
CONTROL CIRCUIT
N
HYBRID
L
RVL′
RVL
$V_{R1}$
(a)
$V_{T1}$
TVL
TVL′
CONTROL CIRCUIT
TVL″
N
HYBRID
L
RVL
$V_{R2}$
(b)
TVL
$V_{T2}$
RVL′
CONTROL CIRCUIT
N
HYBRID
L
RVL″
RVL
$V_{R1}$
(c)
TVL
$V_{T2}$
RVL′
CONTROL CIRCUIT
N
HYBRID
L
TVL′
RVL
$V_{R2}$
(d)

Fig. 10 — Some refinements of the submaster.

### 7.2 *Submaster #2*

A block diagram of this type of speakerphone, with $V_{T1}$ and $V_{R2}$ as inputs to the control circuit, was shown in Fig. 2(b). It has the advantage that even a control circuit with the simple characteristic of Fig. 9(a), without built-in hysteresis, will give satisfactory performance. The reason is that the receive variolosser is located between the incoming line voltage and the control circuit input voltage $V_{R2}$. For a given microphone voltage, the necessary line voltage required to switch the circuit from the transmit to the receive condition is larger than that required to hold it in the receive condition by the amount of loss inserted by the receive variolosser. This is illustrated in Fig. 7(b), where the horizontal shift in decibels can be taken to be exactly equal to the loss of the receive variolosser.

However, to obtain the performance associated with a partially vertical and partially horizontal shift of the transition line like the one in Fig. 7(d), the control circuit must be more complicated and have a characteristic of the type shown in Fig. 9(c). The threshold must be effectively lowered when the circuit is in the transmit condition; a simple arrangement that will boost the gain of the control circuit when transmitting will give the desired characteristic. One way to obtain it is shown in Fig. 10(b), where TVL′ and TVL″ are two extra variolossers, one in each of the two inputs to the control circuit. Both these variolossers should be switched by the control circuit at the same time and in the same direction as the transmit variolosser. The amount of loss switched is about the same in the two variolossers and determines the vertical shift of the transition line.

### 7.3 *Submaster #3*

The inputs to the control circuit for this submaster are $V_{T2}$ and $V_{R1}$, as shown in the block diagram of Fig. 2(c). This type of speakerphone could be considered analogous to the submaster #2 in the sense that a control circuit with a simple characteristic like the one shown in Fig. 9(a) will also result in two transition lines shifted by the right amount in the switching diagram. However, in this case the shift is vertical, as in Fig. 7(c), and equals the amount of loss switched in the transmit variolosser. Receive blocking due to noise operation is likely to occur. As in submaster #2, a better performance could be obtained with a more complicated control circuit characteristic. In this case the threshold should be raised when the circuit is in the transmit condition, as in Fig. 9(d). To accomplish this purpose, Fig. 10(c) shows an arrangement with two extra variolossers, RVL′ and RVL″, which should switch an equal amount of loss. They are operated at the same time and in the same direction as the receive variolosser.

### 7.4 *Submaster #4*

In the interest of completeness, the block diagram of this submaster is shown in Fig. 2(d), with the inputs to the control circuit being $V_{T2}$ and $V_{R2}$. A simple characteristic like the one shown in Fig. 9(a) will give two transition lines in the switching diagram, shifted vertically and horizontally by the amount of loss switched in the transmit and in the receive variolossers. This represents too large an amount of hysteresis. Therefore, as in the submaster #1, the circuit will have blocking problems unless the control circuit characteristic is more complicated. The desired characteristic is shown in Fig. 9(e) and is analogous to that shown in Fig. 9(b) for the submaster #1. Again, a way of obtaining the desired characteristic is to add two extra variolossers, as shown in Fig. 10(d).

## VIII. TRANSIENTS

### 8.1 *Definition of Time Constants*

Since transients concern the time intervals during which the signals are rapidly varying, the time constants associated with the action of the control circuit are very important factors in determining the transient performance. As previously pointed out, two inputs are necessary to determine the action of the control circuit, a voltage from the transmitting branch and a voltage from the receiving branch. It is convenient to define, for each input, two time constants related to the interval necessary for a variation in the input voltage to be detected by the control circuit. Since the control circuit is made sensitive to some weighted average of the input voltages, this process of averaging generally involves a delay. The delay associated with the detection of an increase in the input voltage will be called *build-up time* and will, in general, be different from the delay associated with the detection of a decrease in the input voltage, which will be called *decay time*. There will be, therefore, a build-up and a decay time, $B_T$ and $D_T$, associated with the transmit input and analogous time constants, $B_R$ and $D_R$, associated with the receive input. This is shown graphically in Fig. 11, where, for simplicity, the input voltage is made to vary as a step function and the build-up and decay times are illustrated as the time constants of an exponential curve representing the voltage detected in the control circuit.

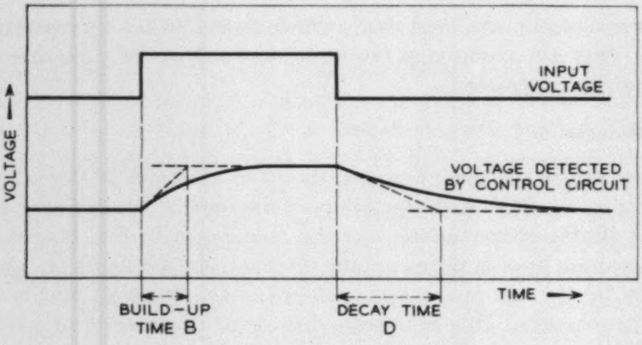

Fig. 11 — Illustration of time constants.

Two more time constants can be defined that correspond to the actual time it takes for the control circuit to switch into transmit or into receive after having detected the two inputs. They will be called $A$ and $R$, the attack and release times of the switch.

It will be assumed in the following that the defined constants are independent, and that the speakerphone is stable in the receive condition when no signals are applied. The latter assumption is in agreement with previous statements about the desirability of having a threshold in the transition line of the switching diagram, as shown in Fig. 7. The delay due to the sound travel time from loudspeaker to microphone and room reverberation will be considered later.

### 8.2 *Relations Among the Time Constants*

Some general observations can be made about the effect of the six time constants on the performance of the speakerphone. When the microphone and the incoming line voltages are varying rapidly, the voltages being detected by the control circuit are affected by the build-up and decay times previously mentioned. Therefore, at each instant one can define effective microphone and line voltages that, under steady-state conditions, would cause the same voltages to be detected by the control circuit as are detected during these transients. The effective voltages are not uniquely defined, since their difference from the actual values depends on how fast the latter ones are varying and on the previous history.

Fig. 12(a) shows qualitatively the shape assumed in some typical cases by the sidetone and acoustic coupling lines when they are plotted in terms of the effective microphone and line voltages. As an example,

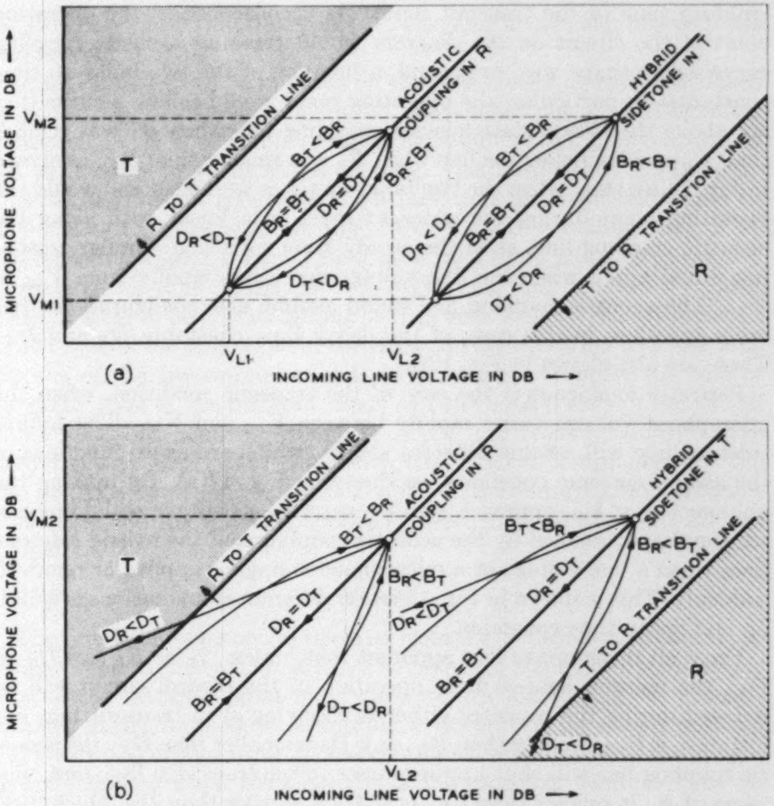
MICROPHONE VOLTAGE IN DB ⟶
$V_{M2}$
$V_{M1}$
T
R TO T TRANSITION LINE
ACOUSTIC COUPLING IN R
HYBRID SIDETONE IN T
$B_T<B_R$
$B_R=B_T$
$D_R=D_T$
$B_R<B_T$
$D_R<D_T$
$D_T<D_R$
T TO R TRANSITION LINE
R
$V_{L1}$
$V_{L2}$
(a)
INCOMING LINE VOLTAGE IN DB ⟶
MICROPHONE VOLTAGE IN DB ⟶
$V_{M2}$
T
R TO T TRANSITION LINE
ACOUSTIC COUPLING IN R
HYBRID SIDETONE IN T
$B_T<B_R$
$B_R<B_T$
$D_R<D_T$
$D_R=D_T$
$B_R=B_T$
$D_T<D_R$
T TO R TRANSITION LINE
R
$V_{L2}$
(b)
INCOMING LINE VOLTAGE IN DB ⟶

Fig. 12 — Effect of time constants during transients.

let us assume that the circuit is in the receive condition and that an incoming line voltage varies between the values $V_{L1}$ and $V_{L2}$. Under steady-state or quasi-steady-state conditions, the operating point of the speakerphone on the switching diagram would move along the straight acoustic coupling line and the circuit would remain in the receive condition all the time. If, however, the line voltage increased very rapidly from $V_{L1}$ to $V_{L2}$, the effective line voltage would follow the variation, with a certain delay determined by the build-up time of the receive input, and would not reach $V_{L2}$ until some time later. Also, the effective microphone voltage, which represents the acoustic coupling, would follow the increase with a certain delay, determined in this case by the build-up time of the transmit input. As a consequence, the operating point of the circuit on the diagram would trace an acoustic coupling curve whose shape was, in general, a function of the two build-up time constants. In particular, the operating point would follow a curve that lay above the steady-state acoustic coupling line when $B_T$ was smaller than $B_R$ and lay below the line when $B_R$ was smaller than $B_T$, as shown in Fig. 12(a). Only when the two build-up times were identical would the operating point during the transient move in a linear path along the acoustic coupling line, as in the steady-state condition. Similar reasoning would apply when the line voltage decreased rapidly from $V_{L2}$ to $V_{L1}$. The acoustic coupling line would assume new configurations that are a function, in this case, of the decay time constants $D_T$ and $D_R$. These are also shown in Fig. 12(a).

Perfectly analogous is the case of the transmit condition, when the microphone voltage varies rapidly between $V_{M1}$ and $V_{M2}$. The hybrid sidetone line will assume different shapes, which are again functions of the same four time constants, as shown in Fig. 12(a). By making the voltage $V_{L1}$ or $V_{M1}$ approach zero ($-\infty$ on a logarithmic scale) one can see the shape assumed by the acoustic coupling and the hybrid sidetone lines when a line voltage or a microphone voltage is applied or removed suddenly. This is shown in Fig. 12(b) for different combinations of build-up and decay time constants.

From all these curves it is apparent that, unless $B_T = B_R$ and $D_T = D_R$, the margins against false operation of the control circuit will be reduced during the transient either in receiving or in transmitting. For instance, if $B_T$ is smaller than $B_R$, or if $D_R$ is smaller than $D_T$, the acoustic coupling line will shift upward closer to the transition line; and, vice versa, if $B_R$ is smaller than $B_T$, or if $D_T$ is smaller than $D_R$, the hybrid sidetone line will shift downward toward the transition line. If the transition line is crossed during a transient, clipping of the signal is likely to occur. Therefore, it is desirable, in general, to satisfy the relations $B_T = B_R$ and $D_T = D_R$.

However, it should be noted that these relations are not very critical. First of all, even with a ratio of 1:2 in the build-up or the decay times and a sudden variation of 20 db in one of the inputs, the maximum departure of the operating point during the transient from the steady-state line would only be about 6.5 db. A smaller or slower variation would tend to reduce this amount. Secondly, even though the transition line might be crossed by the acoustic coupling or the sidetone lines, if this happens for only a very short interval the control circuit will not have time to switch, due to the finite attack and release times, $A$ and $R$, of the switch. Strictly from this point of view, long attack and release times help prevent clipping.

Other conditions to be imposed on the time constants can be deduced by considering a switching operation. As pointed out at the beginning of Section IV, it is desirable to have the circuit switch into transmit as fast as possible to avoid initial clipping. This means that the build-up time of the transmit input and the attack time of the switch should be very short. Experience shows that, with a total switching time of 10 milliseconds, clipping is hardly noticeable, and there is little to be gained by making this time shorter. How to divide this time between $B_T$ and $A$ does not seem to be critical as long as their sum is short. Since the circuit should remain for some time in the transmitting condition to avoid final clipping of the transmitted speech, it is desirable to make the sum of the decay time of the transmit input and the release time of the switch rather long, but not long enough to impair the break-in ability of the distant party. Experience has shown that a good compromise for the sum of these two time constants $D_T$ and $R$ is around 150 milliseconds. This value and the way it is divided between the two constants are not critical.

### 8.3 *Effects of the Acoustics of the Room*

The time it takes for sound to travel from the loudspeaker to the microphone (a few milliseconds) adds to the values of the build-up and decay time of the transmit input only when we are considering the acoustic coupling during a transient. According to the criteria previously outlined and illustrated in Fig. 12, an increase in $B_T$ actually improves the margin against clipping while the circuit is in the receiving condition. An increase in $D_T$ would tend to decrease the margin, but the few milliseconds added are generally negligible compared with the value of $D_T$. Since the hybrid sidetone line or the other time constants are not affected, there is hardly any problem introduced by this effect.

The reverberation in a room, on the other hand, is the cause of the "echo" problem previously described. It has the effect of making the microphone voltage due to acoustic coupling decay at a slow rate after an incoming line signal has ceased suddenly. The reverberation time constant, $R_V$, therefore combines with $D_T$ in determining the shape of the acoustic coupling line. If, for simplicity, we at first neglect $D_T$ compared with $R_V$, the acoustic coupling line assumes one of the forms shown in Fig. 13. These are similar to the lines shown in Fig. 12(b) for different ratios between $D_T$ and $D_R$. Point A is the operating point for a steady line voltage $V_L$. As this voltage is suddenly removed, the microphone

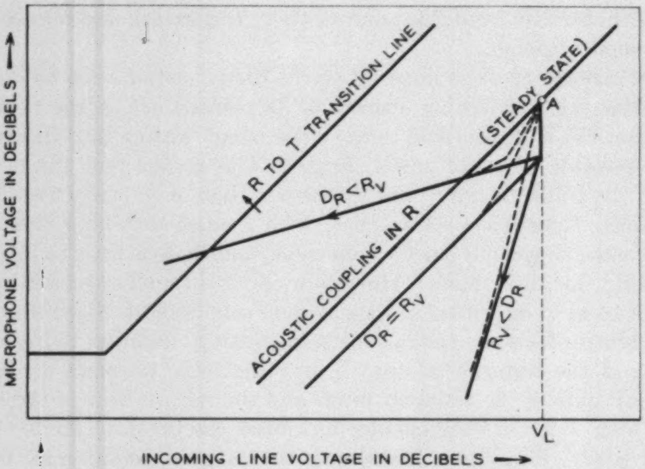

Fig. 13 — Effect of room reverberation.

voltage shows a fast drop, due to the abrupt interruption (within a few milliseconds) of the direct energy coming from the loudspeaker. Then the operating point moves along a straight line whose slope is a function of the ratio between $R_V$ and $D_R$. The larger this ratio, the greater is the chance for the acoustic coupling line to cross the transition line and cause "echo". Therefore, to avoid this problem $D_R$ should be designed to be about equal to $R_V$.

If $D_T$ is not neglected, the curves on Fig. 13 will be slightly different and will be determined by, among other factors, the way the circuit combines the two time constants $D_T$ and $R_V$. In general, the initial drop will not be so abrupt, and the shape will be more like that of the dashed lines. The same conclusions will still be valid.

The reverberation time may vary over quite a range, depending on the size and acoustic treatment of the room, and it may be necessary to adjust the value of $D_R$ at each installation. Experience has shown, however, that, by choosing a fairly long $D_R$, one can take care of the echo problem in most of the rooms encountered in practice without further adjustment. Of course, $D_T$ must also be made relatively long to avoid final clipping of the transmitted signal in a room with short reverberation time. This follows from previous considerations, pointing to the desirability of making the two decay times $D_R$ and $D_T$ about equal.

The general relations derived for the time constants in these sections can be summarized as follows: the build-up and decay times of the transmit input must be approximately equal respectively to the build-up and decay times of the receive input. The sum of the build-up time and the attack time of the switch must be relatively short (about 10 milliseconds), while the sum of the decay time and the release time must be relatively long (about 150 milliseconds). The decay time must be long enough in each case to take care of room reverberation.

### 8.4 *Other Considerations Affecting the Transients*

Important characteristics that affect the performance during the transients are the loss–time curves of the two variolossers. These are the curves that give the variation of the loss with time during the switching interval. To insure proper margin against singing at any instant, the sum of the two losses introduced by the transmit and receive variolossers must be at least equal to the total loss in the steady-state condition. This means that, if the two loss–time curves are plotted on a diagram like the one of Fig. 14(a), the curve for the transmit variolosser TVL must lie at least slightly above the one for the receive variolosser RVL, as indicated by the solid and dashed lines.

That this condition be satisfied is particularly important when the volume control is of the type described in Section VI, which changes the receiving level and the loss switched in TVL and RVL at the same time and by the same amount. A convenient way of arranging this is to have the volume control operate directly on the variolossers, shifting their steady-state receiving condition as shown in Fig. 14(b). As the volume is turned down, a fixed (nonswitchable) loss is introduced in RVL in the receive condition, thus reducing the receiving level by this amount. At the same time, an approximately equal amount of fixed loss is taken out of TVL. The requirements of this volume control arrangement are met if the two loss–time curves satisfy the condition of minimum singing margin at all times.

Besides the relative position of these two curves, their shape is also of interest, because it has a bearing on the unnatural effects of voice switching. To minimize these effects it is desirable to make the rate at which the loss is switched vary in inverse fashion with the amount of energy being changed. For example, in the transmitting variolosser the rate of loss removal in going from the receive into the transmit condition can be rapid at the start of the switching interval, but it should be slower in the final stages when the signals put out on the line are greater. Similarly, in going from the transmit into the receive condition, the restoration of loss should be at a slow rate in the initial stages so that back-

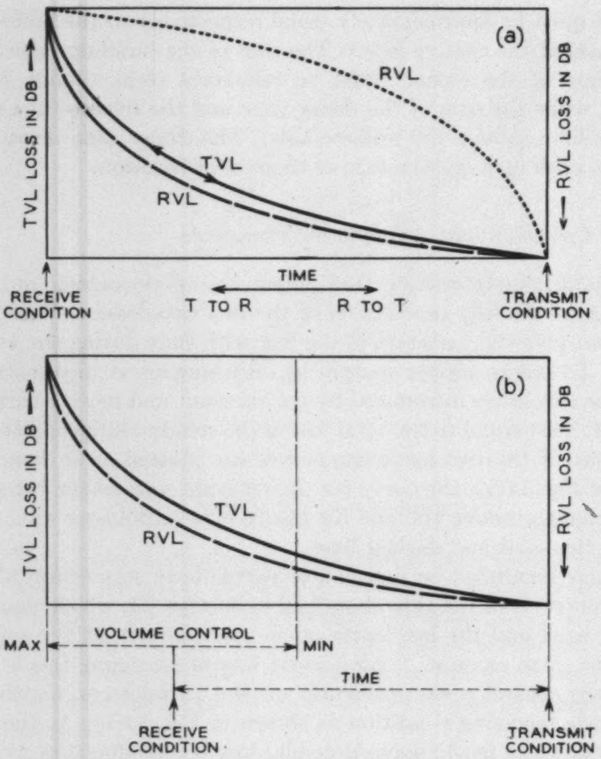

Fig. 14 — Loss-time characteristics of variolossers.

ground noise is not chopped off abruptly after each talk spurt, but a faster rate could be used in the final stages. The resulting shape is the one shown in Fig. 14(a) for the transmit variolosser.

If the same criteria were applied to the receive variolosser, the RVL loss-time curve should have the shape of the dotted line in Fig. 14(a). It is obvious that this RVL curve is not compatible with the chosen TVL curve if the singing margin must be maintained at all times. However, it is more important to minimize the unnatural effects of voice switching in transmitting than in receiving, since the circuit tends to remain stable in receiving when no signals are applied. On this basis, the loss-time curves for the transmit and receive variolossers should assume the shapes shown in Fig. 14(b).

## IX. ANALYSIS OF SOME SPECIAL FEATURES

A problem commonly encountered in the use of speakerphones is sensitivity to local noises, whether they are sharp loud noises like the slamming of a door or the steady type like that of a fan. In the case of sharp noises, if $B_T$ and $A$ are short and $D_T$ and $R$ are long, as was shown to be desirable, the noise will switch the circuit into transmit and keep it there for some interval, interfering with the reception of the incoming signal. One way to reduce this interference is to design the control circuit so that $D_T$ and $R$ are short for sounds of very short duration and have the normal long value only when a sound persists for a period equal to the shortest duration of speech sounds. This "deferred hangover" principle has been successfully used in TASI.[8]

With steady noise, the problem is to keep the circuit from switching into transmit and causing receive blocking. A simple solution consists in raising the threshold of the transition line. However, in order not to impair the performance of the speakerphone in a less noisy environment, the threshold should be adjusted in each case to be just above the noise level. There are ingenious circuits that make this adjustment automatically. In order to operate properly, they must discriminate between speech and steady noise, so that the threshold will not be affected by the former but will be raised by the latter. The fluctuating characteristic of speech energy compared with noise is used for this discrimination.

## X. SUMMARY

In the design of a speakerphone circuit and in the evaluation of its performance from both the steady-state and the transient viewpoints, the switching diagram described in this paper has proved a useful tool. The fundamental problems of voice switching have been classified, and criteria for avoidance of them have been given in terms of the diagram characteristics and the circuit time constants.

Four submasters representing general configurations of a speakerphone with only the essential features, have been examined. Each of them can be made to give comparable performance, and the choice in general will depend on circuit design considerations, mainly on which type results in the simplest and most economic design for a given application.

At the cost of some increase in complexity, features can be added to the submasters to give improved performance in a particular characteristic. Some of the ingenious features that have been used or previously suggested have been examined. In general, these can be analyzed and evaluated by means of the switching diagram.

## REFERENCES

1. Clemency, W. F., Romanow, F. F. and Rose, A. F., The Bell System Speakerphone, Comm. & Electronics, No. 30, May 1957, p. 148.
2. Emling, J. W., General Aspects of Hands-Free Telephony, Comm. and Electronics, No. 30, May 1957, p. 201.
3. Wright, S. B. and Mitchell, D., Two-Way Radio Telephone Circuits, B.S.T.J., **11**, July 1932, p. 368.
4. Hopkins, H. F. and Stryker, N. R., A Loudness Rating for Loudspeakers and Power Requirements for Rooms, Proc. I.R.E., **36**, March 1948, p. 315.
5. Wright, S. B., The Vodas, B.S.T.J., **16**, October 1937, p. 456.
6. Tillman, R. J., U.S. Patent No. 2,269,565.
7. Laurence, R. F., A Voice-Switched Loudspeaking Telephone for Public Network Lines, British Comm. and Electronics, February 1957, **4**, p. 91.
8. Bullington, K. and Fraser, J. M., Engineering Aspects of TASI, B.S.T.J., **38**, May 1959, p. 353.

# Integrated Magnetic Circuits for Synchronous Sequential Logic Machines

By U. F. GIANOLA

(Manuscript received November 16, 1959)

*The magnetic switching core is a small, reliable and inexpensive component for digital logic circuits, but conventional core circuits also include a large number of auxiliary semiconductor components. As a result, the cost and reliability of core circuits are not very different from those built entirely with semiconductors. The possibility of building logical systems that use a minimum of nonmagnetic components is examined. Except for the provision of clock-pulse sources, fully integrated magnetic machines are feasible in principle. Suitable circuits must contain provision for gain, memory and unilateral transmission of data. These requirements can be met by taking advantage of the threshold characteristics of ferromagnetic materials that have a rectangular hysteresis loop. Two approaches are selected for consideration. In the first, a binary digit is represented by the remanent state of magnetization of a ferrite core. Upon the application of clock pulses its state can be transferred to adjacent cores by means of electrical interconnections. In the second approach, a binary digit is represented by a discrete flux pattern in a continuous flux conductor. These patterns can be propagated in a step-by-step process through the flux conductor by means of clock pulses. Examples are given of experimental synchronous sequential circuits using commercially available multiapertured cores. The capabilities, limitations and organization of these circuits are discussed.*

## I. INTRODUCTION

The ferrite memory core has many attractive features as a switching element for digital computer circuits.[1,2,3,4,5,6] It combines a memory function with the circuit equivalent of a switch, and it is small and inexpensive. Furthermore, as a magnetic amplifier it can provide gain. It is also believed to be reliable, the only ambients to which it is susceptible being temperature, external magnetic fields and mechanical stress. In spite of this, its use in logical circuits has been mostly limited to access circuitry for magnetic memories.

Why is this? One reason is that the maximum rate at which present-day cores can be cycled is around 1 mc, the limitation being the rise in temperature produced by hysteresis losses. However, this speed is sufficient for many systems, and materials are being continually improved in this respect. A more pertinent reason is to be found in the circuit approaches that have been used. Additional circuit components, and diodes and transistors in particular, are usually included. The result has been that the total count of semiconductor components in some cases — for example, shift registers — is not materially reduced compared to the equivalent all-semiconductor circuits. Thus, the reliability and cost are set mostly by the semiconductor components. In addition, in order to provide a correct impedance match between the magnetic and semiconductor components, it is necessary to use multiturn windings on the cores, and this considerably increases the cost of assembly. Therefore, the more general use of magnetic core logic is dependent upon the design of new circuits that materially reduce or eliminate semiconductor components. This article describes the possibilities that have been examined, and it will be shown that an almost complete integration in terms of magnetic components may be possible. Means rather than applications will be stressed.

Section II is tutorial in nature, and is intended to introduce the reader to the basic tenets of magnetic core circuitry including the terminology and models used, and also to the requirements of a fully integrated circuit. In Section III the requirements of combinational and sequential logic machines are discussed, and possible approaches to the design of a magnetic machine are introduced. An approach based upon the use of multiapertured cores is described in detail in Section IV, and experimental data are presented. Section V describes approaches based upon the step-by-step propagation of flux patterns through a continuous flux conductor; there is little experimental information available in this case. Finally, the state of the art is summarized in Section VI. A full treatment of this nature may help elucidate the possibilities and limitations of this particular approach to an integrated computer technology. It is hoped that dividing the article into sections in the manner described will simplify the selection of points of interest.

## II. GENERAL PROPERTIES OF CORE CIRCUITS

For completeness, an account will be given of the properties of memory cores and the requirements of an integrated circuit.

We shall be concerned with ferromagnetic materials that have the rectangular hysteresis characteristic shown in Fig. 1.[7,8] Note the sharp threshold for switching from an upper to a lower state of magnetization,

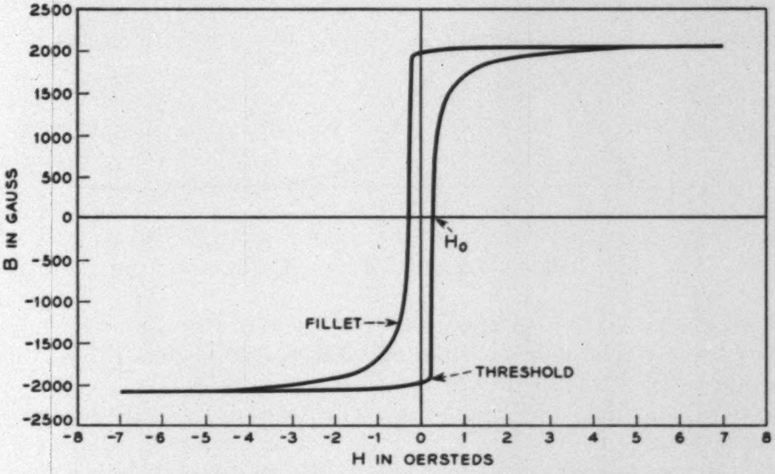

Fig. 1 — *B–H* characteristic of a typical fast switching memory core ferrite.

or vice versa. An objectionable feature is the fillet, which is found in all of the ferrite materials and to a lesser degree in the permalloys, since it means that the switching threshold must be exceeded by an appreciable factor if a complete reversal is to be obtained. Sometimes it is perfectly satisfactory to operate on a minor loop, in which case the fillet is not troublesome, but for many of the circuits to be considered it would be better if the fillet were eliminated. It is to be hoped that materials will be developed to accomplish this.

In the majority of cases to be considered, lines of magnetic flux can be completed entirely within the material, and the effect of air leakage can be neglected. This means that the magnetic field $H$, in oersteds, produced by an applied drive $NI$, in e.m.u., is given approximately by the expression $H \cong 4\pi NI/L$, where $L$, in centimeters, is the mean peripheral length of the core, irrespective of the manner in which the winding is placed on the core. This approximation may be invalid in cases where portions of the magnetic path are saturated. If a constant-current drive is used to reverse the flux, the rate of change of flux as a function of time, $d\varphi/dt$, has the form shown in Fig. 2. The switching time, $\tau$, is normally defined as the separation between the 10 per cent points of the $\dot{\varphi}$ versus $t$ waveform, and is approximately inversely proportional to the applied switching field, as is shown in Fig. 3.[8] The resultant total flux reversal,

$$\int_{\tau} \dot{\varphi}\, dt,$$

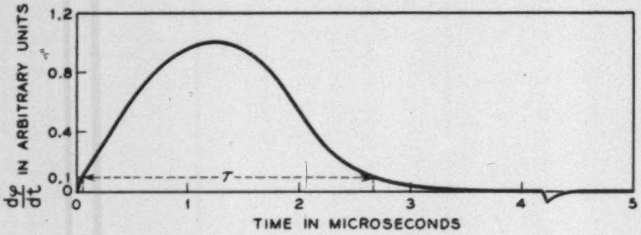

Fig. 2 — $d\varphi/dt$ as a function of time.

varies with drive in the manner shown in Fig. 4. At large drives the total flux reversal approaches a constant value $\Phi$. The gradual rise to this value can be attributed to the fillet.

It is customary when analyzing core circuits to make the following approximations:[2] (a) $d\varphi/dt$ is constant during switching; (b) $\tau(H - H_0) =$ constant $= s_w$; (c) the material has an ideal rectangular loop. In other words, it is assumed that it is possible to replace the waveform of Fig. 2 by one that is rectangular with the same enclosed area and time duration $\tau$; that of Fig. 3 by a straight line having the same mean slope and axis crossing; and that of Fig. 4 by a step function having the same mean amplitude $\Phi$ and axis crossing $H_0$. These assumptions lead to the useful approximation for large drives

$$\frac{d\varphi}{dt} \cong \frac{\Phi}{\tau} \cong \frac{\Phi(H - H_0)}{s_w} \cong \frac{4\pi\Phi}{s_w L}(NI - I_0),$$

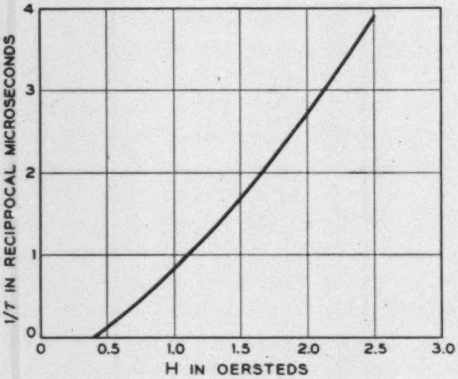

Fig. 3 — Reciprocal switching time as a function of applied field.

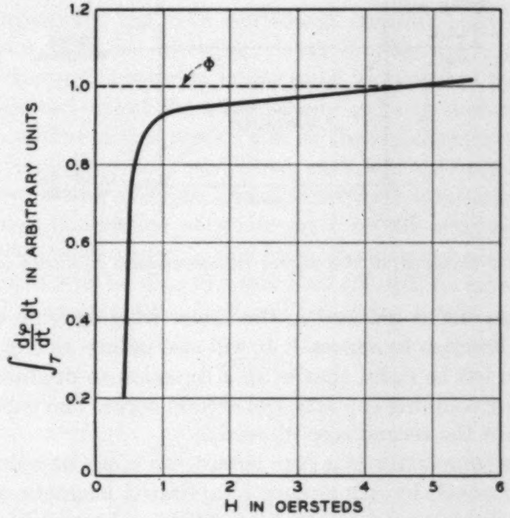

Fig. 4 — Total flux reversal as a function of applied field.

where $NI$ is the resultant constant-current drive and $I_0$ is the threshold current for switching using a single-turn winding. Units are in e.m.u. Typical values for existing ferrites are $s_w = 0.7 \times 10^{-6}$ oersted-second* and $\Phi = 4 \times 10^3 \times A$ gauss-cm², where $A$ is the minimum cross section of the core. The model is a crude one, but it has been found to give quantitative values useful for designing circuits that use either simple toroidal cores or multiapertured cores.

It is customary to describe magnetic circuits schematically by means of mirror symbols.[2] Toroidal cores are represented by heavy vertical lines, which may be regarded as end view projections. Winding leads are represented by horizontal lines and the winding turns coupled to the cores are indicated by short diagonal line segments ("mirrors") at the intersections. A simple example is given in Fig. 5. The sense of the field produced in the core is obtained by "reflecting" the applied current at the corresponding mirror surface. Thus, in Fig. 5, $I$ will set the first core downwards. The sense of the voltages produced in coupled windings by a flux change are obtained by reversing the applied field and "reflecting" it at each mirror surface. Thus, in Fig. 5 as the core is set

---

* Switching time may also be defined as the time interval between 10 per cent Φ and 90 per cent Φ points. For this definition, $s_w \cong 0.2 \times 10^{-6}$ oersted-second. The relevance of either definition will vary with the application.

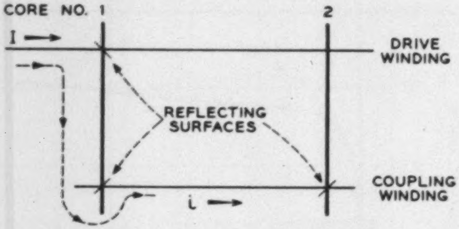

Fig. 5 — Example of the mirror representation of a core circuit.

downwards an emf is induced in the upper winding from right to left, that is, in a direction to oppose $I$. It will also induce an emf in the lower winding from left to right, that is, in a direction to produce a current $i$ in the winding coupling the first and second cores. The current $i$ is in a direction to set the second core upwards.

The wiring complexity of a core circuit can often be reduced, and its efficiency increased, by using more complicated magnetic structures in place of the simple toroidal cores.[9,10,11,12,13] In order to follow the operation of these structures, it is only necessary to keep in mind two facts. First, flux continuity must be preserved. Second, when an mmf exceeding threshold is applied, a flux reversal will first take place in the switching path presenting the lowest reluctance in preference to the remaining paths.[10] With few exceptions, the lowest reluctance path is the shortest path having flux available for the reversal.

Most multiapertured cores can be reduced to one of two types, or to a combination of the two. Representative of the first is the laddic structure,[10] which has the property that all important flux paths are flux-limited to the same extent. Representative of the second type is the transfluxor,[9] which has the flux capacity of one path equal to the sum of the flux capacities of the remaining paths. In a sense, the laddic can be considered as consisting of a number of toroidal cores interconnected in series by magnetic linkages, while in the transfluxor the toroidal cores are interconnected in parallel. The magnetic linkages make possible a transfer of flux from one core to another without the use of a coupling winding. The latter, because of added resistive losses, reduces the efficiency of circuits.

A modified mirror symbol representation will be used to describe circuits made with multiapertured cores. The procedure is illustrated in Fig. 6 for a two-hole transfluxor. Each leg of the core shown in Fig. 6(a) is represented by a heavy vertical line in Fig. 6(b), the notation at the bottom giving the flux capacity set by the minimum cross section of the corresponding leg. The horizontal striated lines are included simply to draw attention to the fact that the legs are connected magnetically, so that a flux reversal in one must be balanced by an opposite reversal in one or both of the other two in order to maintain flux continuity. The preferred flux return will be through the closest available path. Apart from these last restrictions, each vertical may be treated as a separate core. As an example, assume that leg 1 is initially set up, legs 2 and 3 down. Application of a current *I* to the winding shown will drive flux downwards in leg 1, and maintain downwards saturation in leg 3. The result will be that one flux unit (Φ) will be reversed in leg 1 and returned through leg 2, which will be left set in the upwards direc-

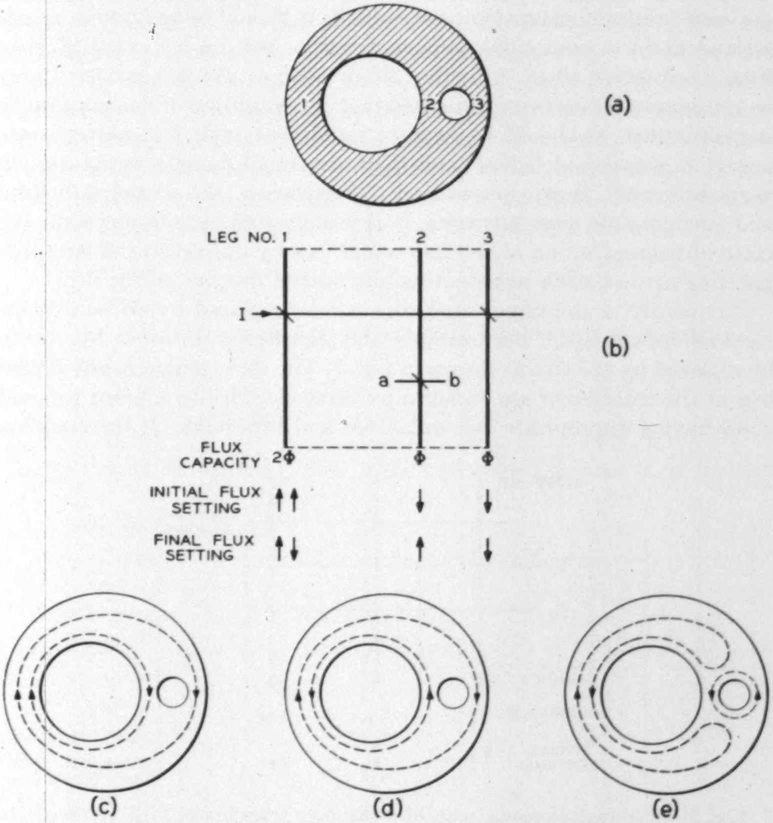

Fig. 6 — Transfluxor, its mirror-symbol equivalent, and flux patterns.

tion. Applying the normal mirror circuit procedure, it will be seen that the result of setting flux up in leg 2 will be to induce an emf in the winding *ab* in a direction to drive current from terminal *a* through the winding to terminal *b*. This notation simplifies the explanation of the more complicated circuits.

It is often helpful in describing the operation of these structures to represent the state of magnetization graphically by means of closed flux patterns. As an example, three of the possible states of magnetization of the transfluxor are shown in Figs. 6(c), (d) and (e). In this case, each of the flux arrows represents the saturation remanent flux capacity of each of the smaller legs. Thus, in Fig. 6(c) all legs are fully saturated; whereas in Figs. 6(d) and (e) legs 2 and 3 only are saturated, leg 1 having zero resultant magnetization—which, it should be remarked, is not arrived at by a sinuosidal demagnetization, but by a partial reversal from a saturated state. A further subdivision of the flux patterns may be introduced if necessary, provided that the condition of flux continuity is maintained. It should be kept in mind that such flux patterns are merely a convenient fiction, and that the actual domain structure will be considerably more complicated. Nevertheless, the model is helpful and adequate for most purposes. It is usually sufficient to represent the state of magnetization of the individual legs by upwards or downwards pointing arrows, each representing one unit of flux, as in Fig. 6(b).

Frequently, a multiapertured core can be replaced by an equivalent toroidal core circuit.[13] For example, the transfluxor shown in Fig. 6 can be replaced by the circuit shown in Fig. 7. The three magnetically linked legs of the transfluxor are replaced by three electrically coupled toroidal cores having appropriate flux capacities and thresholds. If the coupling

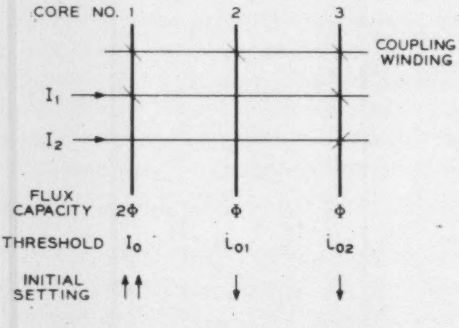

Fig. 7 — Equivalent core circuit of a two-hole transfluxor; $I_0 > i_{02} > i_{01}$ , in order to allow for the difference in switching path lengths between the corresponding legs of the transfluxor.

winding has zero resistance, then flux may be transferred from core 1 to core 2 by $I_1$, and then from core 2 to core 3 by a current-limited drive $I_2$, as between the corresponding legs of the transfluxor. In general, however, the operation of the multiapertured core equivalent is simpler to explain. It is also inherently more efficient, because one part of the advance sequence is carried out by a flux transfer within the core itself, so that for this flux transfer there is no power dissipated in coupling windings. Furthermore, the use of multiapertured cores reduces the cost of assembly, since the cost of putting in coupling windings appears to be the most important single factor.

It is possible to obtain a flux gain between stages of a core circuit by using an adequate turns ratio in the coupling winding. In order to illustrate this point, consider the circuit shown in Fig. 8(a). Two cores are coupled by a closed winding of resistance $R$ having $n$ turns on core 1, but only one turn on core 2. Let core 1 be initially set up and core 2 set down by $I_1$. Now apply $I_2$ to reverse the flux in core 1. Simultaneously a current $i$ will be induced in the coupling winding in a direction to reverse the flux in core 2. The distributed inductance and capacitance of the winding can be neglected in a practical instance. Thus, considering instantaneous electromotive forces in the coupling winding,

$$iR = n\frac{d\varphi_1}{dt} - \frac{d\varphi_2}{dt},$$

where the subscripts 1 and 2 denote core 1 and core 2, respectively. Accordingly,

$$\int_0^t d\varphi_2 = n\int_0^t d\varphi_1 - \int_0^t iR\,dt.$$

Clearly, when $R \to 0$ the flux change produced in core 2, if its flux

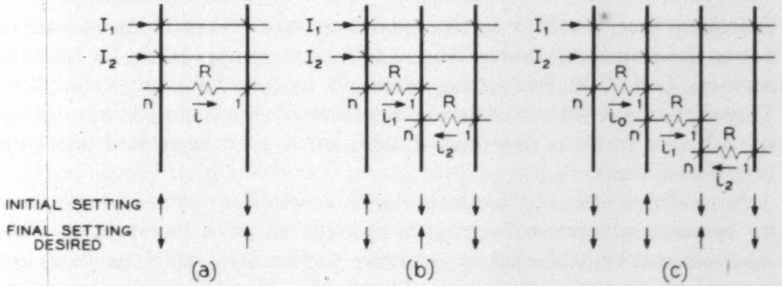

Fig. 8 — Core circuits to demonstrate the capability for gain and the need for decoupling.

capacity is sufficient, will be $n$ times larger than that produced in core 1, and, for a finite value of $R$, it may be made at least equally large by a suitable choice of $n$.

However, the provision of a turns ratio is not in itself always sufficient for providing gain. To illustrate this point, consider the circuit shown in Fig. 8(b). Cores 1 and 2 are coupled together as previously, but in this case core 2 is also coupled to a following core by a similar winding. Core 1 is initially set up and cores 2 and 3 set down by $I_1$. Core 1 is now driven down by $I_2$ in order to transfer its setting to core 2, as previously. However, in this case, the resultant mmf acting on core 2 is reduced, because an opposing current $i_2$ is induced in the winding coupling core 2 to core 3. Considering instantaneous electromotive forces in the coupling windings:

$$i_1R = n\frac{d\varphi_1}{dt} - \frac{d\varphi_2}{dt},$$

$$i_2R = n\frac{d\varphi_2}{dt}.$$

Note that there is no $\dot{\varphi}_3$ term, because there is no flux reversal in core 3. Furthermore, according to the approximate model for the core circuit described previously,

$$\frac{d\varphi_2}{dt} \cong \frac{\Phi}{\tau_2} \cong \frac{4\pi\Phi}{s_wL}\,(i_1 - ni_2 - I_0).$$

Therefore, introducing the values of $i_1$ and $i_2$ in terms of $\dot{\varphi}_1$ and $\dot{\varphi}_2$,

$$\frac{d\varphi_2}{dt} = \frac{n\dfrac{d\varphi_1}{dt} - I_0R}{1 + n^2 + \dfrac{s_wLR}{4\pi\Phi}}.$$

It follows that, for $R > 0$, the maximum rate of switching of core 2 is less than one-half that of core 1. Thus, it is impossible to transfer even one-half of the flux setting of core 1 to core 2 for any value of $n$. Therefore, some method of decoupling forward coupling windings during the advance phase is necessary in addition to a turns ratio if flux gain is to be obtained.

An additional reason for including a decoupling feature is the need for preventing adverse backwards propagation of a flux setting. This situation can be explained by reference to Fig. 8(c), which again shows a series of interconnected cores. In this case flux is initially set up in core 1 and down in the remainder by $I_1$. During the advance phase the flux in core 1 is to be driven downwards by $I_2$, in order to advance its initial setting to core 2, as previously. However, in this case a current $i_0$ is also induced in the back winding coupling core 1 to core 0. This current is in a direction to set flux upwards in core 1, and therefore there must also be provision for preventing this.

In conventional core circuitry the necessary decoupling is achieved by including a diode or transistor in the coupling windings. It is its function that must be replaced by magnetic components in a fully integrated circuit.

The question of decoupling and gain has been dealt with at some length because its understanding is crucial to the design of integrated magnetic circuitry. The need for gain to compensate for circuit losses and to permit fanout is obvious, but the failure to recognize the need for decoupling is a common pitfall. The analysis presented here is based upon a relatively crude model of a core circuit, and the behavior of an actual circuit may be somewhat worse or possibly better, depending upon the actual properties of the cores. It may be better if the core has a re-entrant hysteresis characteristic of the kind found in some of the permalloys.[14] In these materials once the static threshold ($H_0$) is exceeded and switching commenced, the threshold for continued switching decreases as shown in Fig. 9. This can result in an additional flux gain relative to that calculated from a model based upon the ideal rectangular characteristic. However, this possibility has not been studied extensively and so will not be considered further.

## III. REQUIREMENTS OF COMBINATIONAL AND SEQUENTIAL CIRCUITS

Simple logic machines, which must generate a unique Boolean function of a set of input variables that are coincident in time, are best constructed with combinational circuits for fastest performance. For these systems the situation is fairly satisfactory. For example, any switching function can be generated by a single laddic or a combination of other magnetic components. Semiconductor components may be required in the input and output circuits in many cases, but the total count relative to completely semiconductor circuits is small for complicated switching functions.

More complicated machines, particularly general-purpose machines, must be constructed with sequential circuits. The logical inputs in this case are presented sequentially in time, and the outputs obtained at any particular time are used to provide the inputs during the next phase of the sequence. In other words, the state of magnetization of a magnetic circuit set during a first phase must be capable of determining the setting

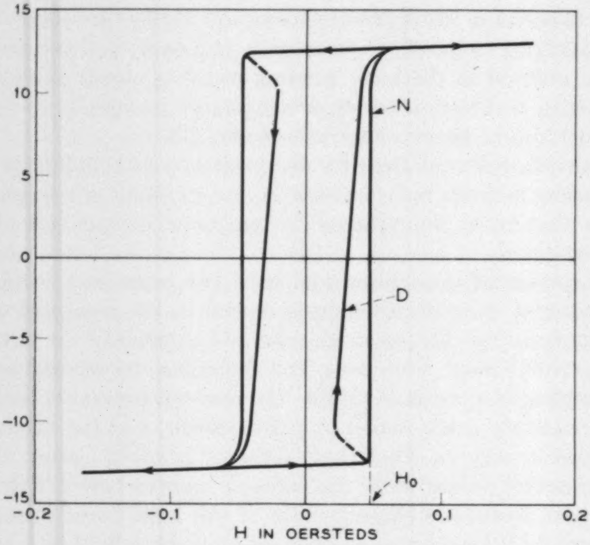

Fig. 9 — Example of a re-entrant hysteresis loop found in stressed 69-permalloy. In order to initiate a flux reversal a field exceeding the static threshold, value $H_0$, must be applied. However, once switching has commenced it can continue under an applied field less than $H_0$. The normal loop (N) is measured in the conventional way. The dynamic loop (D) is obtained by first applying a field exceeding $H_0$ to initiate switching and then immediately reducing it.

of another during a second phase. There must, therefore, be provision for storage, gain and unidirectional transfer of a sequence of bits. The remainder of this paper will be mainly concerned with this kind of organization.

There are two basic approaches possible. The first, which is considered in Section V, is more in the nature of a true magnetic circuit concept in that it involves the step-by-step propagation of discrete flux patterns through a continuous magnetic structure. It is the more elegant approach because the bit information may be kept in magnetic form throughout an entire sequence of logical operations, being translated to electrical form only at the input and output terminals.

The second approach is more conventional; separate cores are used for each combinatorial phase of the operation and are interconnected by coupling windings. The core itself provides the required storage function, and, as discussed in Section II, a flux gain between stages can be obtained by means of a suitable turns ratio in the coupling windings, provided that there is also provision for decoupling at the appropriate times and for preventing backwards propagation of the bit information. As mentioned previously, in conventional core circuits the decoupling is obtained by including a diode or transistor to block undesired currents. It is the function of these diodes that must be reproduced. It should be noted that only the nonlinear resistance property of the diode is important. Other nonlinear components could be used, but the present discussion will be limited to consideration of the use of the nonlinear property of the cores themselves. The fact that a current drive must exceed a certain amplitude before it can produce a flux reversal means that the core can be considered to have a current sensitive switch in series with its input. The function of the diode could be replaced by such a switch. All circuits proposed so far using both simple toroidal cores and multiapertured cores make use of this feature.[15,16,17] As mentioned in Section II, multiapertured core circuits have definite advantages. For this reason circuits using simple toroidal cores will not be considered. Rather, in Section IV the discussion will be limited to a consideration of a multiapertured core circuit that fulfills the basic requirements. Because a certain amount of flux steering can be accomplished within the multiapertured cores, their use actually represents a compromise between the two approaches.

For small systems it is possible to perform a limited amount of sequential logic using laddics and related devices as the circuit elements. These possibilities have been described previously.[10] It should also be noted that parametron and ferroresonant magnetic core circuits also fulfill the stated requirements.[18,19] Complete logical systems based upon the former have been reported. However, they do not enter naturally into the present discussion, which is concerned primarily with the use of cores exclusively. It should only be remarked that they appear to share similar speed–power limitations.

## IV. CORE CIRCUITS

The problem of organization of a synchronous sequential machine is a complex one. For the present class of circuits, the most profitable starting point is to use a shift register having associated combinational components. As noted previously, all the operations for combinational logic at a given stage of the advance sequence can be performed on a single core, which might, for example, take the form of a laddic with a transfluxor output. This paper will mostly be concerned with the design of the shift register.

The circuits to be described in this section are based upon the trans-

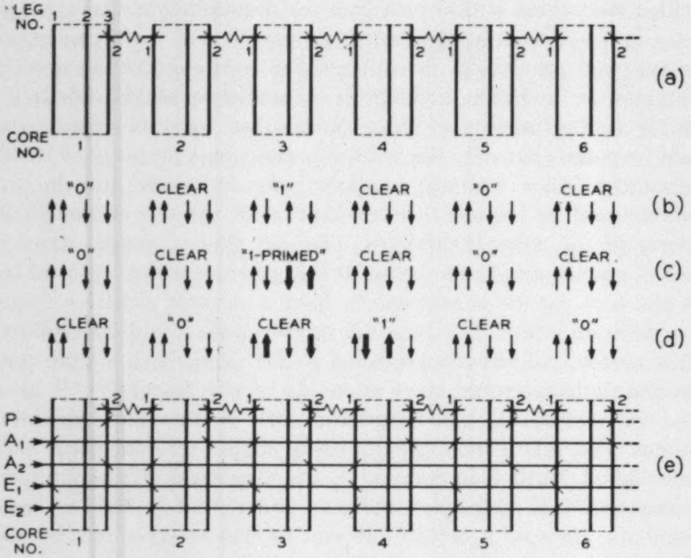

Fig. 10 — (a) Skeleton shift register circuit; (b) initial flux setting; (c) primed flux setting; (d) advanced flux setting; (e) complete circuit.

fluxor structure shown in Fig. 6. The framework of the shift register is shown in Fig. 10(a). Leg 3 of each core is coupled by a closed winding to leg 1 of the following one. The resistance shown in Fig. 10(a) is the resistance of the winding itself, normally a few tenths of one ohm. A turns ratio is used to give the flux gain necessary to make up for transfer losses.* In practice, it is convenient to make $n = 2$. The flux setting of each leg will be represented by arrows in Figs. 10(b), (c) and (d), each arrow representing one unit of flux ($\Phi$).

Alternate cores are used as information cores at each stage in the advance sequence; the remainder are in a "cleared" state. In other words, at one stage in the sequence the bit information will be stored in the odd-numbered cores, at the next stage it will be advanced into the adjacent even-numbered cores, etc. As an example, in Fig. 10(b) the odd-numbered cores 1 and 5 are shown set in the "zero" state, core 3 being set in the "one" state. The corresponding flux patterns are shown in Fig. 6(c) and Fig. 6(d) respectively. These two states correspond to the "blocked" and "unblocked" states of a transfluxor.[9] The "clear" state also corresponds to the "blocked" condition. Observe that it has been possible to change the flux pattern to insert a one into core 3 without producing a flux change in leg 3, which would induce an emf in the winding coupling core 3 to core 4. Therefore, it was possible to effect a full flux transfer from core 2 to core 3 without including the diode, which was required in the conventional toroidal-core circuit discussed in connection with Fig. 8(b).

---

* It should be possible to use a unity turns ratio in the coupling loop by taking advantage of the flux gain mechanism produced by a re-entrant $B$–$H$ characteristic as mentioned in Section II. However, the materials property necessary is somewhat critical, and as a result the alternative complication of a turns ratio is preferred at present. Furthermore, the flux gain is then sufficiently large that the allowable tolerances are improved.

The bit information may be transferred from the odd-numbered cores to the following even-numbered cores, leaving the former in the "clear" state, as follows:

The information cores are first "primed" by applying a drive to leg 3 of every core in a direction to switch flux upwards. This is the *P* (prime) drive shown in the complete circuit, Fig. 10(e). It is limited in current amplitude so that although it can switch flux between legs 2 and 3, that is, around the small hole of the transfluxor, it cannot produce a flux reversal around the longer path including legs 3 and 1. The *P* drive will not affect the zero or "clear" cores. However, in a core which contains a one it will set leg 3 up and leg 2 down. In other words, the flux pattern is changed from that shown in Fig. 6(d) to that shown in Fig. 6(e). As leg 3 is being set up, a current will be induced in the winding coupling cores 3 and 4. However, the direction of this current is such that it does not change the flux setting of core 4. The final setting of all cores following the *P* phase is shown in Fig. 10(c). Note that it was at this point that advantage was taken of the threshold characteristic of the core, in this case to prevent the improper conversion of a zero into a one. The closest corresponding situation in the conventional toroidal-core circuit was the adverse backwards propagation of a one.

Observe that the current that was induced in the winding coupling cores 3 and 4 was limited only by the winding resistance. This current acted in opposition to the applied drive, and therefore reduced the mmf acting on leg 3. It follows that leg 3 switched quite slowly during this phase, and that a full reversal will have taken place only if the *P* current was applied for a sufficient length of time. It is not necessary for the successful operation of this circuit to produce a complete reversal, because attenuation of flux by an incomplete reversal can be compensated for by the gain provided by the turns ratio of the coupling loop. However, the flux reversed must exceed a certain minimum, and so there is a corresponding minimum time for bit propagation.

The information cores are now to be "cleared" and the bits advanced. To do this, a drive is applied to the $A_1$ (advance) winding shown in Fig. 10(e). Cores 1 and 5 are already "blocked", so that there is no change in their flux patterns nor in those of the coupled cores 2 and 6. On the other hand, core 3 is driven from the "one-primed" to the "clear" state. This flux reversal induces a current in the coupling winding between cores 3 and 4 in a direction to reverse flux in leg 1 of core 4. The $A_1$ drive prevents a flux reversal in leg 3 of either core 4 or core 2. Therefore, the sum result is to set leg 2 of core 4 in the upwards direction to give the flux pattern shown in Fig. 10(d). The bit information has now been advanced into the even-numbered cores. Subsequently, the latter can be primed (by $P$), and then cleared (by $A_2$) to advance the bits one further stage back into the odd-numbered cores. In this manner by applying the sequence of drives $P$, $A_1$, $P$, $A_2$, the bits may be propagated from left to right through the register.

This circuit may be analyzed by means of the approximations described in Section II. In the Appendix, it is shown that the resistance, $R$, of the coupling winding has a maximum permissible value because of the requirement that there be a flux gain of at least unity between stages. The minimum switching time, $\tau_p$, possible for a $P$ phase is obtained by making the winding resistance equal to this maximum value. For a large ratio of the diameter of the large hole to that of the small hole $(\tau_p)_{\min}$ is approximately proportional to $s_w/H_0$. Since $s_w$ cannot be reduced appreciably below existing practical levels, it is necessary to increase $H_0$ in order to decrease $\tau_p$. Therefore, higher switching speeds can be obtained only at the expense of a corresponding increase in drive power.

It is not necessary for the resistance of the coupling winding to be exactly equal to the maximum value in order to obtain the maximum speed for the $P$ phase, since, as noted previously, a partial flux reversal is permissible. Therefore, the time duration of the $P$ phase may be set by the width of the drive pulse rather than by the actual switching time, permitting considerable latitude for the choice of the resistance of the winding. If speed is unimportant, there is no lower limit for $R$, since, given sufficient time, a flux reversal during the $P$ phase can always be completed irrespective of the resistance of the coupling winding.

The switching time during an $A$ phase is limited only by the amplitude of the current drive practicable. Therefore, the speed of the circuit is limited primarily by the $P$ phase, and the cycle time required to advance a bit one stage is approximately equal to $2\tau_p$. Measured values of the minimum value of $\tau_p$ for three commercial multiapertured cores are given in Table I. It will be seen that the 3665 core is capable of a bit handling capacity, $\frac{1}{2}(\tau_p + \tau_A)^{-1}$, of the order of $5 \times 10^4$ per second. By improving the design of the core as discussed, it should be practical to increase this figure by perhaps an order of magnitude, but a further increase in speed will probably depend upon the development of improved materials, since the drive requirements become excessive.

The peak power requirement occurs during the $A$ phases, when a flux reversal must be produced around the large hole of each core containing a one. At the same time, current must be supplied to make up for that taken by the coupling windings. The total drive current required using single turn windings is approximately $I_t n^2/(n - 1)$, where $I_t$ is the current required to produce a flux reversal around the large hole of an unloaded core in time $t$. Thus, for $n = 2$, the advance current is $4I_t$. The peak back emf developed in a single turn for each bit is approximately equal to $\Phi/t$. Therefore, the peak power requirement is equal to $4I_t \times \Phi/t$, or approximately $s_w \Phi L/\pi t^2$. In a typical situation, this peak power is approximately 0.4 watt/bit for $t \sim 1$ microsecond. The average power needed is less by a factor $t/T$, where $T$ is the total cycle time for advancing a bit. The power required for the $P$ phase is less, since switching takes place at a slower rate. The average power dissipation in watts/bit is approximately $10^{-6}$ times the bit rate for the commercial cores. It is interesting to note that almost exactly the same figure has been reported for parametron and ferroresonant shift register circuits.[18,19]

TABLE I—MINIMUM SWITCHING TIMES AND OPERATING CURRENTS FOR SHIFT REGISTERS USING THREE COMMERCIAL CORES

| Core number | $E_1$, ampere-turns | $E_2$, ampere-turns | $P_1 = P_2 = P$ (max.), ampere-turns | $P_1 = P_2 = P$ (min.), ampere-turns | $A_1 = A_2$, ampere-turns | $(\tau_p)_{\min}$, microseconds |
|---|---|---|---|---|---|---|
| RCA XF-3665 | 0 | 0 | 0.56 | 0.46 | 9 | 8 |
| | 0.5 | 0 | 0.56 | 0.46 | 4 | |
| RCA XF-3668 | 0 | 0 | 0.39 | 0.27 | 5 | 30 |
| General Ceramic F-1023-S6 | 0 | 0 | 0.14 | 0.12 | 2.4 | 38 |
| | 0 | 0 | 0.15 | 0.10 | 3 | |
| | 0.1 | 0 | 0.14 | 0.12 | 1.2 | |
| | 0 | 0.1 | 0.21 | 0.12 | 3 | |

The $A$ drive current has wide margins. It is only necessary that it be sufficient to clear the cores within a few microseconds. On the other hand, the magnitude of the $P$ drive currents must be kept within fairly well defined boundaries, as shown by Fig. 11. Curve A shows how the total flux switched around the small hole of an "unblocked" core varies with the applied drive, and curve B shows the same relation for switching around the large hole of the core when "blocked." The *P* drive must lie within the boundaries delineated by the hatched region. If it exceeds the right-hand boundary it will produce unwanted switching in a core containing a zero, so that zeroes may gradually be converted into ones. On the other hand, if it falls below the left-hand boundary insufficient flux will be switched in a core containing a one, so that the one will become attenuated even though there is provision for a small flux gain between stages. The boundaries shown in Fig. 11 are the measured values obtained from experimental re-entrant registers. They can be considerably improved by suitable design of the geometry of the core.

Additional drive windings can be incorporated to increase margins, to increase speed and to reduce the drive current amplitudes. For ex-

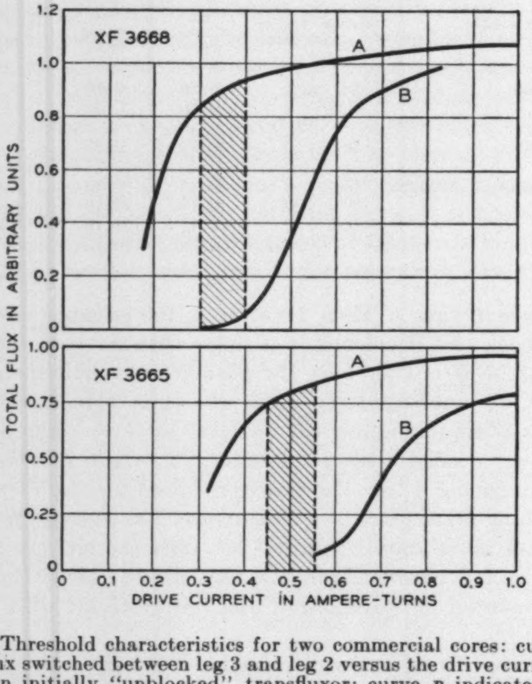

Fig. 11 — Threshold characteristics for two commercial cores: curve A indicates total flux switched between leg 3 and leg 2 versus the drive current applied to leg 3 of an initially "unblocked" transfluxor; curve B indicates total flux switched between leg 3 and leg 1 versus the drive current applied to leg 3 of an initially "blocked" transfluxor.

ample, in Fig. 10(e) an additional winding, $E_1$, is shown. It may be used during each $A$ phase to bias cores, which are receiving bit information, up to the threshold needed to switch flux between legs 1 and 2. It therefore reduces the current required in a coupling winding to switch a one into a core and, therefore, the magnitude of the $A$ drive necessary to complete switching in a prescribed time. A second winding, $E_2$, can be used during the $P$ phase. It provides an mmf around the large hole that acts counter to the $P$ drive, making possible an increase in the upper limit of the $P$ drive and the corresponding maximum bit rate. Other windings of this nature can be incorporated. However, all such windings can at best only increase the speed by a factor of two. These additional windings can normally be driven in series with the corresponding major drive windings, so that additional current sources are unnecessary. In this connection also, the $P$ drive can conveniently be supplied by a dc source, so that only two pulse sources are necessary, one for $A_1$ and another for $A_2$. Thus, in a sense, only a two-phase drive circuit is needed, but there must be adequate spacing between $A_1$ and $A_2$ pulses to allow for the slow priming operation, due to the use of a dc source for the $P$ current.

The circuit can be modified in various ways to reduce the peak current amplitudes required. As an example, consider the circuit shown in Fig. 12(a). It uses a core that has small holes between legs 1 and 2 and between legs 3 and 4, and a large hole between legs 2 and 3. Each leg in this case has a flux capacity, $\Phi$, of one unit. The additional small hole makes it possible to clear the information cores in the back direction during the $P$ phases, thereby eliminating the additional current that

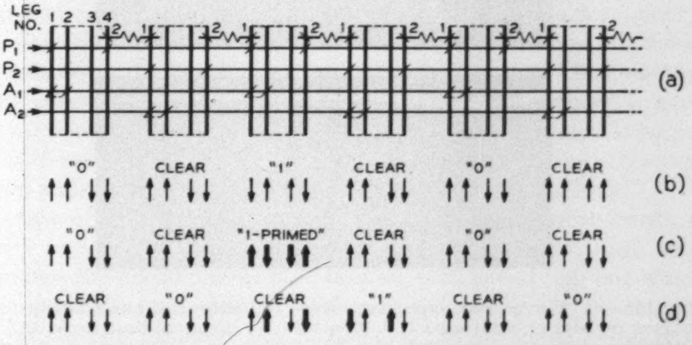

Fig. 12 — (a) Modified circuit; (b) initial flux setting; (c) primed flux setting; (d) advanced flux setting.

the $A$ drive must provide in the circuit of Fig. 10(e) to compensate for that dissipated in the coupling loop in the back direction. The operation of this circuit can be followed by studying the sequence of flux patterns shown in Figs. 12(b), (c) and (d). The cycling pattern in this case is $P_1$, $A_1$, $P_2$, $A_2$, $P_1$, $\cdots$. A dc source can again be used for both $P_1$ and $P_2$.

Experimental shift registers have been built based on both the circuits of Figs. (10(e) and 12(a) using each of the cores listed in Table I. They have been operated in re-entrant circuits, i.e., with the output of the final core in the chain coupled directly to the input of the first without intermediate amplification. Fig. 13(a) shows the output that was obtained from one of the cores in a re-entrant three-bit chain using XF3665 cores in the circuit of Fig. 10(e) when a one, zero, zero was be-

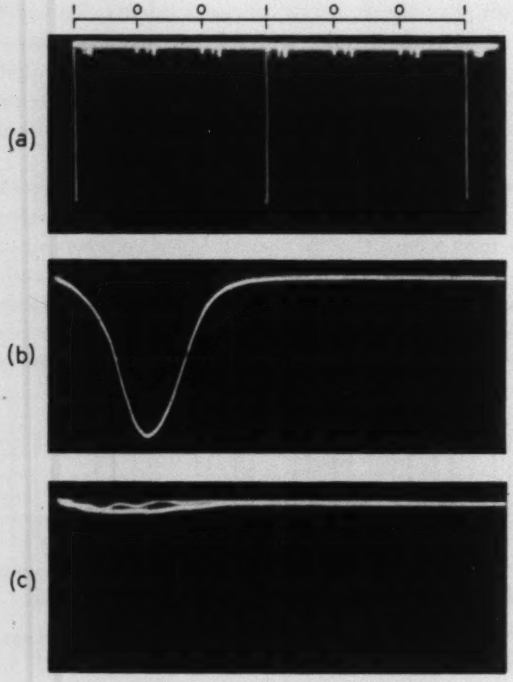

Fig. 13 — (a) Waveforms appearing across the output of one transfluxor in a re-entrant register circulating a 1, 0, 0 pattern. Vertical calibration = 0.15 volt/turn/division; horizontal calibration = 500 μs/division. (b) Output waveform for a one. Vertical calibration = 0.15 volt/turn/division; horizontal calibration = 0.1 μs/division. (c) Output waveform for a zero. The maximum noise signal is also shown superimposed. The calibration is the same as that for (b).

ing continuously circulated. Figs. 13(b) and (c) show the corresponding waveforms for a one and a zero on an expanded time base. The operating currents are shown in Table I. The values given for $A_1(=A_2)$ are the minimum values needed for reliable operation using the corresponding minimum values for $P_1(=P_2)$.

The ampere-turns required to drive these circuits are large, but not beyond the capabilities of transistors. For example, the 2N417 and the 2N580 can be used to give pulses between 1 and 2 amperes in amplitude. In addition, components such as the General Electric controlled-rectifier C35F can provide pulses up to 150 amperes peak amplitude. They have been used as switches to discharge pulse-forming network drivers. Relay contacts have also been used. The waveform of the drive pulses is not critical. A rise time of 0.5 microsecond is adequate.

It should also be remarked that it is a simple matter to introduce multiturn drive windings in order to reduce the current for a given ampere-turn drive. The cores can be mounted in such a manner that each turn consists of a single wire threaded straight through an in-line assembly of alternate cores. Multiple turns can then be obtained by interconnecting the individual wires of a multiconductor cable at the ends of the register. A photograph of an experimental 10-bit register module is shown in Fig. 14. The cores were assembled and wired upon the two jig rods shown in (a), and then pushed together for close packing. The circuit was then mounted on a printed circuit card and potted, as shown in (b). It was unnecessary to preselect the XF3665 and F-1023-S6 cores used in re-entrant registers, except to reject those that were obviously mechanically damaged. However, this was not the case for the XF3668 cores. Registers up to 800 bits (1600 cores) capacity have been operated.[20]

The register can be interrogated nondestructively at each stage of the shift, since the flux patterns for a one and a zero correspond to those of a transfluxor in the "unblocked" and "blocked" states. Thus, each core can be interrogated nondestructively in the normal manner.[9] It is convenient to use an additional small hole in the structure for this purpose. Fig. 15 shows the waveforms obtained upon nondestructively interrogating two register cores containing a one and a zero respectively.

Because of the gain feature, fanout is practical. In other words, one or more inputs may be driven in parallel by each output. Each fanout will reduce the maximum bit rate, because each added coupling winding will increase the total counter mmf produced during a $P$ phase by the induced currents. If there are $N$ fanouts, the switching speed during a $P$ phase is reduced by a factor a little less than $N$. The $A$ drive on the core must also be increased by a corresponding factor in order to

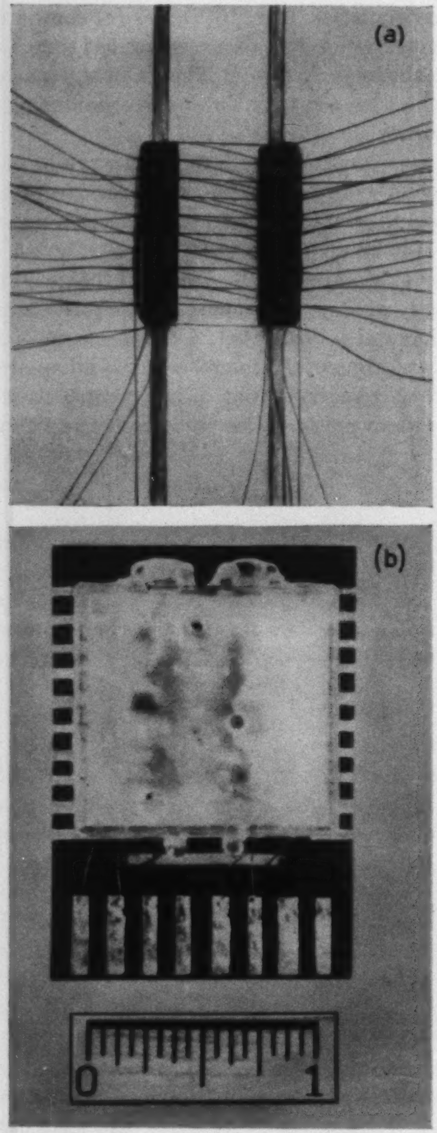

Fig. 14 — (a) Ten-bit shift register cores assembled on jig. There are ten parallel inputs on the left-hand side, and ten parallel outputs on the right-hand side. The clock windings run parallel to the jig rods. (b) Register mounted on printed circuit card and potted.

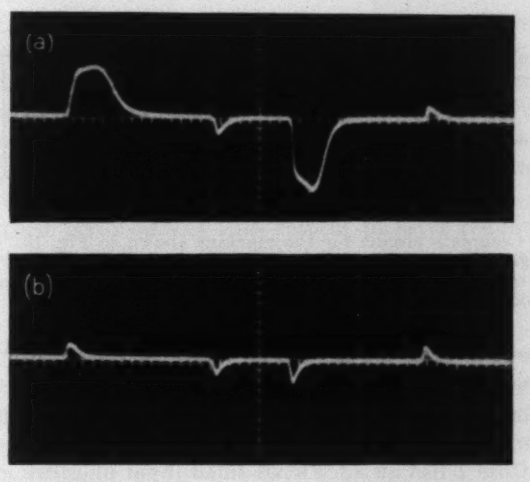

Fig. 15 — (a) Output waveform obtained by nondestructive interrogation of a core containing a one. (b) Output waveform obtained when the core contained a zero. Vertical calibration = 0.1 volt per division; horizontal calibration = 2 $\mu s$ per division.

compensate for the increased total of coupling currents loading the core. Experimentally, a fanout of three has been demonstrated.

As an example of the use of fanout and the incorporation of an additional simple combinational operation, consider the circuit shown schematically in Fig. 16, and in its entirety in Fig. 17.[21] Outputs $x$ and $y$ are inserted into an exclusive-OR circuit. The output of the latter provides the input for the register. It can easily be shown that this circuit will successively generate all combinations of four bits except four zeroes. For if the register initially contains 1000, during the next advance the exclusive-OR circuit will insert a zero into the input of the register, which will now contain a 0100; subsequent advances will give 0010, 1001, etc., for a total of 15 terms before the initial setting is repeated.

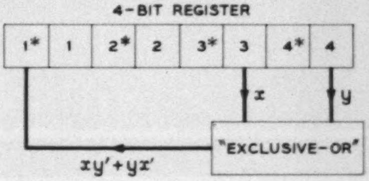

Fig. 16 — Example of a synchronous sequential machine that can be constructed entirely with multiapertured cores.

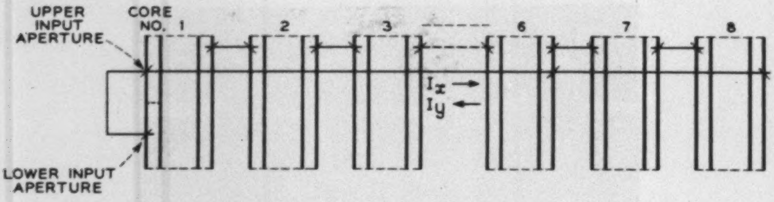

Fig. 17 — Complete circuit for the machine shown in Fig. 16. The *P* and *A* drive windings are omitted for clarification, since they are exactly the same as those shown in Fig. 12, with the exception of core 6, which must have a doubled $A_1$ drive winding because of the fanout at that point.

The three-hole core circuit of Fig. 12(a) was used for the register. The exclusive-OR function was derived by connecting the outputs from the third and fourth bit position of the register, cores 6 and 8 respectively, in series opposition to drive the input core, core 1, as illustrated in Fig. 17. Core 1 has an additional small input aperture, obtained, as indicated in the mirror circuit representation, by segmenting the first aperture with a magnetic interconnection. When $I_x$ alone is present, a one is inserted into core 1 by means of a flux reversal around the path including leg 3 and the lower input section of leg 1. When $I_y$ alone is present, the reversal path includes leg 3 and the upper input section of leg 1. When both $I_x$ and $I_y$ are present, or both are absent, the resultant drive is zero, and there is no switching in core 1, i.e., a zero is inserted. The output from a single core in this circuit shows the repetitive pattern of 15 settings as the bit patterns circulate, Fig. 18.

In summary, it appears that, with sufficient ingenuity, core circuits can be designed that should satisfy the stated objectives. This can be done at the expense of a considerable reduction in the maximum system speed possible compared to a core-diode system using the same cores,

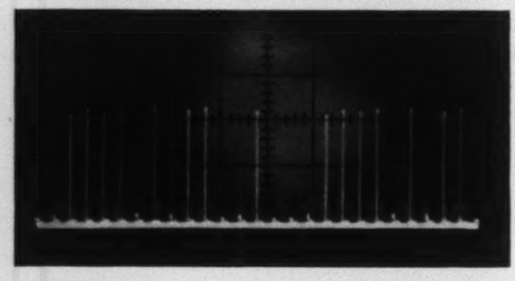

Fig. 18 — Output obtained from an experimental model showing the repetitive pattern of fifteen bits that was generated, i.e., 001111010110010.

but for many applications in the telephone system speed is not a primary requirement. For elementary systems, the operating conditions have been shown to be reasonably practical and can be improved by designing the device with this application in mind.

As mentioned previously, two other circuits based upon the transfluxor have been described in the literature,[16,17] and there are a number of variations possible. Their operation differs in detail, but an analysis shows that their performances with respect to drive requirements and speed are generally similar. A detailed analysis giving the design parameters of these circuits is presented in the Appendix.

## V. SHEET LOGIC

The idea of performing logic by a step-by-step transfer of a flux pattern through a continuous magnetic structure is an appealing one for the reasons given previously. Known methods of this kind, unfortunately are especially reliant upon the threshold characteristics of the materials, and all advance drives must be amplitude-limited. Furthermore, there is no provision for gain. However, the techniques are of interest and there are some potentially useful applications, so that a detailed consideration is justified.

Two complementary approaches will be considered. The first makes use of the threshold characteristics of multiapertured structures,[6,13] while the second adds the influence of the depolarizing fields produced when the structure is open, i.e., when the flux pattern is closed by air-flux outside the structure.[22,23,24] A third possibility is based upon the use of flux summation structures. However, the latter can usually be reduced to toroidal core circuits, which have been described in the literature[6,25,26] and do not appear to fulfill all the requirements for a synchronous sequential machine. Therefore, this approach will not be considered.

An elementary example illustrating the propagation of flux in a multiapertured structure is shown in Fig. 19. A laddic structure is used, its feature being that the flux capacities of the rungs, and preferably that of the side-rails also, are all equal. To begin with, the flux pattern shown in Fig. 19(a) is established. The flux around alternate windows is magnetized in an anticlockwise sense to the left of rung 5, and in a clockwise sense to the right. Rung 5 is left in a neutral state of magnetization. A first advance pulse, $A_1$, is now applied. It produces an mmf in a direction to drive flux downwards in rungs 3, 5, 9, 11, 15, 17, $\cdots$, as shown in Fig. 19(b). Of these, only rung 5 has flux available for a reversal. There are two possible return paths for this reversal: to the left of rung 5 through rung 1, or to the right of rung 5 through rung 7. The first

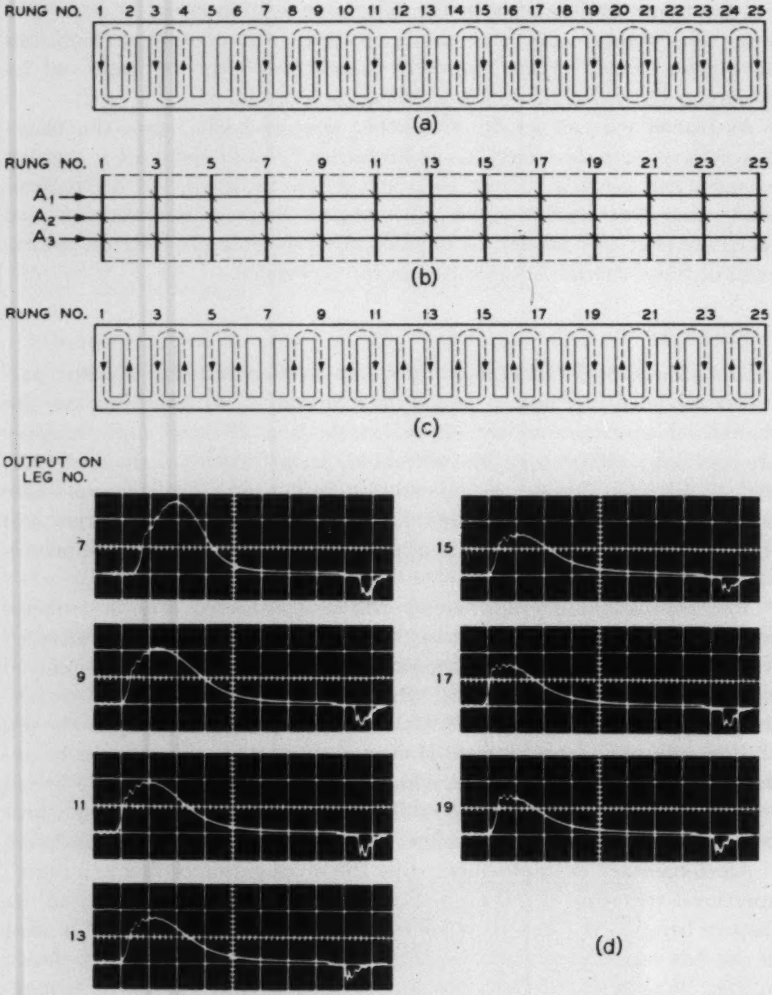
RUNG NO. 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
(a)
RUNG NO. 1 3 5 7 9 11 13 15 17 19 21 23 25
A₁
A₂
A₃
(b)
RUNG NO. 1 3 5 7 9 11 13 15 17 19 21 23 25
(c)
OUTPUT ON LEG NO.
7
9
11
13
15
17
19
(d)

Fig. 19 — (a) Initial flux pattern in laddic stepping switch; (b) circuit diagram; (c) flux pattern following one advance step; (d) successive outputs obtained from an experimental structure as the bit pattern was propagated through successive odd-numbered rungs.

path is almost twice as long as the second. Therefore, if the $A_1$ current is limited, so that it is below the threshold required for switching between rungs 5 and 1, only the reversal between rungs 5 and 7 can take place. Thus, following the $A_1$ drive the flux pattern becomes that shown in Fig. 19(c). It will be seen that the initial flux pattern has been displaced two rungs to the right to rung 7. Similarly, the flux pattern will be displaced an additional two rungs to the right following each of the successive drives $A_2$, $A_3$, $A_1$, $\cdots$ . In this manner, the bit information represented by the neutral magnetization of a rung can be propagated step-by-step from left to right. By adding a fourth phase, it is possible to use the same structure as a self-clearing shift register.

Many alternative arrangements are possible, making use of modified advance windings or structure, or both. However, they are all subject to the same limitations, depending very much on the threshold characteristics of the materials, and require amplitude limited drives throughout. As a result at each step, a little flux may be transferred beyond the required point, (or left behind), because of an inferior threshold characteristic. In addition, any departure from the correct geometry — e.g., an overlarge rung — will also result in a deterioration of the flux transfer. These effects are cumulative, so that the result is a continuous attenuation of the flux pattern. For example, in an experimental model the flux transferred to the correct rung was reduced by approximately 7 per cent at each step. As an example, the successive outputs obtained as the bit pattern was propagated through the successive odd-numbered rungs, 7, 9, 11, 13, 15, 17 and 19 respectively, are shown in Fig. 19(d).

The second method is best introduced by considering the example of a magnetic tape, which has an easy direction of magnetization parallel to its long axis. Let it be fully magnetized in this direction, say from right to left. Now let a field be applied in a direction to reverse the magnetization in a small segment, *ab*, as illustrated in Fig. 20; for example, by means of a small solenoid. In this case, flux closures must be completed by air-flux (not shown). Under these conditions, the segment *ab* is subjected to demagnetizing fields, which increase as the length *ab* is decreased, and, in fact, if *ab* is too small the reversed setting will be unstable once the applied field is removed. In consequence, it will be possible to produce a stable zone of reversed magnetization only if *ab* exceeds a critical length, which is dependent upon the thickness, width and magnetic properties of the tape.

Once a stable reversed segment is produced, it can be extended if a pulsed current is applied to an adjacent solenoid that is less than the critical length. This solenoid would not be capable of establishing a

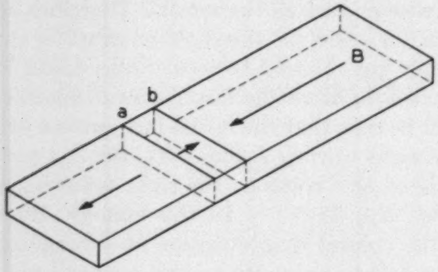

Fig. 20 — A reversed segment in a magnetic tape.

permanent reversal in the absence of an adjacent reversed segment, but it can do so if the latter is present, because the demagnetizing field is reduced. Therefore, it is possible to propagate a bit, which, in this case, is represented by the presence or absence of a reversed segment bounded by a 180° butt domain wall, in a similar manner to that used for propagating the bit pattern in the laddic. The corresponding circuit is shown in Fig. 21, which uses mirror symbols to show the sense of the windings, $A_1$, $A_2$ and $A_3$. Each winding acts upon every third segment, e.g., *ab*, *de*, etc., and each segment is less than the critical length but greater than one-half of it. Assume that the tape is initially magnetized from left to right on the left of *a*, and in the opposite direction on the right. Now apply the $A_1$ drive. It acts to reverse magnetization in every third segment — *ab*, *de*, etc. — but, because each is less than critical in length and alternate segments are spaced by more than one critical length, it will succeed in producing a permanent reversal in *ab* only, thereby advancing the wall from *a* to *b*, as discussed. Similarly, as the successive drives $A_2$, $A_3$, $A_1$, $\cdots$ are applied, the wall is advanced step-by-step from left to right.

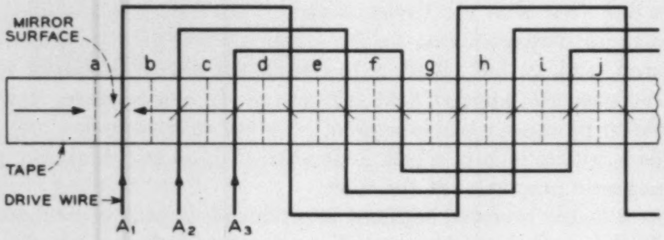

Fig. 21 — Circuit diagram of a magnetic strip stepping switch.

As for the laddic structure, the threshold characteristics of the material must be well defined, and the drives must be amplitude-limited, since the effective length of each segment of the reversal field depends upon both the current amplitude and the length of the solenoid. Both the laddic and tape devices operate in a similar fashion, and synchronous sequential circuits using either can be designed. However, the latter are simpler to describe and also appear to be the more promising from a practical viewpoint. Therefore, they will be used as the basis for further discussion.

The function of the circuit shown in Fig. 21 is that of a stepping switch or counter, since an emf will be induced in turn in output windings placed at intervals along the structure. This must be reset every time the butt wall reaches the terminal output winding. The uses of such a device are limited, but, by a slight modification of the advancing circuits, it is possible to restore the magnetization of the strip to its reset direction to the rear of the reversed segment as the latter is being propagated. In this case, it becomes possible to propagate a series of reversed segments simultaneously, and the device can be used as a shift register.[22,23] A minimum of four advance phases is necessary.

This circuit could be used as the basic building block of a synchronous sequential machine, as discussed for the core circuits. However, if the register must provide parallel outputs for the electrical input to a similar device, the transferring drive current should not be limited to the same extent as are the advance currents. In other words, the transfer drive should be capable of supplying the additional ampere turns required in the output winding, in addition to the drive required for the operation of the unloaded register. This requirement can best be met by designing the circuit so that the output winding is coupled to the tail end of the bit segment at the appropriate clock time. The output emf is then produced by the drive, which serves the primary purpose of restoring the magnetization of the end section of the bit segment to the reset direction, thus clearing the tape to the rear. This drive need not have the same amplitude limitation that was imposed upon the advance currents.

A shift register circuit, which uses electrically interconnected short tapes in much the same way as cores were used in Fig. 10, is illustrated in Fig. 22. A stable reversed segment, *ab*, representing a one in the first tape can be advanced into the adjacent segment, *bc*, which is smaller than the critical length, without modifying the setting of the adjacent tape, which is in the "cleared" or zero state. This comprises the priming operation. The first strip is then reset to the "clear" state (by an unlimited amplitude drive), thereby inducing a current in the coupling winding in a direction to set a one into the following tape in its *ab* segment. Reverse propagation can be inhibited as before. The interesting feature of this circuit as compared to the core circuit is that the magnetic structure is threaded through the windings, rather than vice versa. This could simplify the assembly advantageously.

It will be clear that the strip register meets the requirements of a basic building block for a synchronous sequential machine. Fanout can be achieved by electrical interconnections between tapes, as in the transfluxor circuits. However, in this case an interesting and reasonably practical possibility is to obtain a fanout by an actual physical branching. In practice, this can be achieved by replacing the single tape with a number of separate flux conductors, which are driven in common up to the branch point but individually beyond there, where necessary. Fan-in can be similarly obtained. The total number of branches is limited, since each fanout will reduce the magnitude of the flux quantity representing a bit. However, this is not a serious limitation.

In principle at least, it should be possible to make quite extensive synchronous sequential machines by these techniques. Thin magnetic film or wire technologies appear particularly suitable for their fabrication. Unfortunately, experimental work in this field has been very limited, so that the actual capabilities are unknown. However, there has been sufficient experimentation to demonstrate the feasibility of these proposals.[22,23] The critical length of a stable reversed segment is of the order of one millimeter in nickel-iron films and wire. The actual range and uniformity of the critical length for different materials have not been studied extensively, but an examination of the shape of a reversed segment suggests that there is going to be a practical minimum limit. To

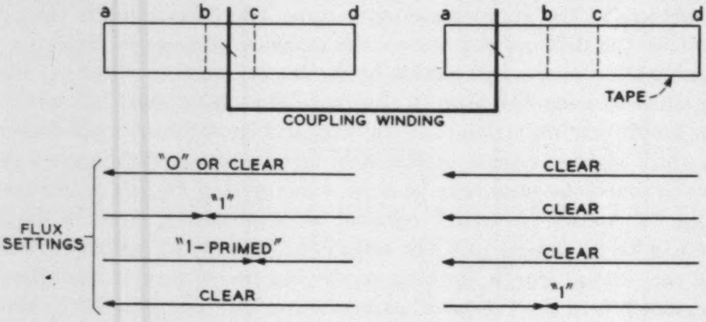

Fig. 22 — A magnetic strip shift register circuit that is analogous to the core circuit of Fig. 10.

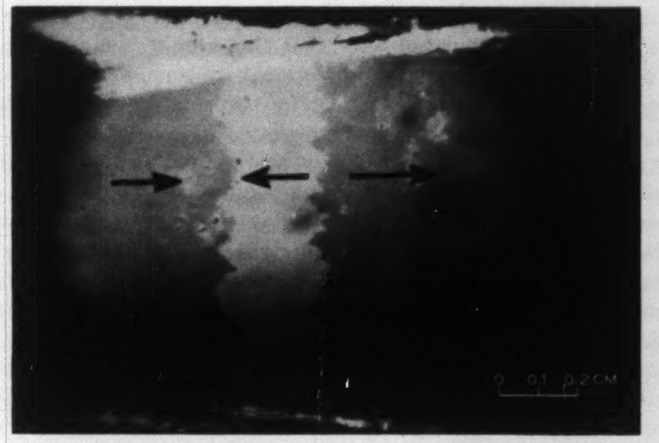

Fig. 23 — Photograph of a 180° butt wall domain.

illustrate this point, a photograph of the structure of a reversed segment in a thin nickel-iron film is shown in Fig. 23. The photograph was taken through a polarizing microscope, which delineated the region of reversed magnetization by the different rotation of the plane of polarization that it produced. It will be seen that the domain wall is very irregular, varying in position by approximately 0.25 mm. The critical length must, of course, be much larger than this irregularity, or its threshold will be poorly defined.

The question of speed is an interesting one. The limiting speed of an unloaded continuous-tape shift register is set by the time required for an unstable reversed segment to revert to normal magnetization when an advancing field is removed. This time will decrease as the length of the segment is decreased, so that the greatest speed is obtained by making each segment equal to its minimum value, which is one-half of the critical length for the four-phase shift register. The reversal time can be decreased by using a material having a higher threshold field, all other parameters being kept unchanged, but at the same time the drive current must be increased. Thus, in this case also, increased speed can only be obtained at the expense of an increased drive.

## VI. DISCUSSION

An attempt has been made to coordinate the known ways to exploit the properties of rectangular loop magnetic materials in order to build synchronous sequential machines using a minimum of other components. The requirements stipulated were that the circuits should have a capacity for gain, memory and unidirectional transmission of data. These requirements were shown to be met by two related systems, the first using an assembly of electrically interconnected cores, and the second using an assembly of continuous-flux conductors; the latter must use electrical interconnections where gain is required, and a modicum of flux steering can reduce winding complexity in the former. Of the two, the first is immediately practical, at least for simple subsystems such as shift registers, and commercially available devices are suitable if system requirements are not excessive. In principle, neither system is limited in speed, but increased speed can be obtained only at the expense of an increase in drive power. It cannot be proved, but it appears at the present time that this limitation applies generally to all magnetic logic systems. Circuits of this nature may have a constant power–speed product analogous to the constant gain–bandwidth product of vacuum tube circuitry.

As matters stand, circuits using electrically interconnected cores appear to be satisfactory when bit-rate requirements are not excessive. They are expected to be reliable and economical. Fully integrated magnetic systems are feasible, in principle, with the exception of a few nonmagnetic components which must be included to provide the primary pulse sources. Such a machine should be designed at the very beginning in terms of magnetic components, rather than attempting to use magnetic subsystems as replacements for parts of an existing design. The replacement approach very often suffers because of the complications afforded by an impedance mismatch, as well as a mismatch of the relative capabilities of different technologies. The balance of parallel versus serial organization of a subsystem also should be kept in mind when speed is a consideration. The apparent reliability and low cost of magnetic devices may make practical an increased use of parallel organization. Questions of this nature will always be somewhat controversial, and indeed can only be settled for quite specific applications.

The question of interrogation of the outputs of a machine has been touched on only briefly. A magnetic device can be interrogated in a number of ways, both destructively and nondestructively. Most of these have been discussed in detail in the literature.

The number of magnetic devices for memory and logic is continually increasing, so that it would not be justified to assume that the devices selected for consideration in this paper represent the optimum. However, it does appear likely that alternatives will make use of similar techniques. No doubt more difficult, and occasionally simpler, approaches will become evident. Few devices more complicated than shift registers have been constructed in this technology, and the practicability of more complex systems has yet to be established. The objective is to construct suitable function packages, which can be interconnected for specific applications. Because of the low input and output impedances, it is desirable to make individual packages considerably more complex in logical performance than has been done in other technologies. For example, in the case of the elementary shift register, let us not ask for a single-bit package to replace those now available, but rather for a multibit package. This is quite practicable; indeed, it is the most economical approach, and gives the most compact assembly. The further development of these circuits may open up some interesting possibilities for computer design.

## VII. ACKNOWLEDGMENTS

I have to thank T. H. Crowley and W. D. Lewis for introducing me to this subject. The former has also cooperated in a number of paper studies. I have also frequently drawn upon the work and support of many others, especially A. H. Bobeck, E. B. Ferrell, D. L. Hofmockel, P. Mallery and H. J. Schulte. E. M. Walters assisted in all of the experimental work, and G. Knapp kindly supplied the photograph of a butt wall domain. R. A. Chegwidden, F. R. Monforte and W. W. Rhodes have been very cooperative in providing ferrite materials, and the group headed by J. F. Muller in providing workshop facilities.

## APPENDIX

The limiting parameters of the circuit described in Section IV can be determined by means of the approximate model for a core discussed in Section II.

First, to determine the properties required of the coupling winding consider the part of the shift register circuit shown in Fig. 24, let core 1

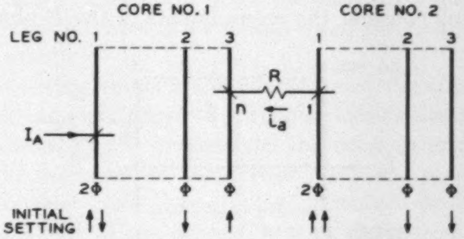

Fig. 24 — Shift register circuit, *A* phase.

contain a "one-primed" setting which is to be advanced into the second core as the $A$ drive clears the first. In order that the maximum amount of flux be transferred, core 2 must switch at least as fast as core 1 at this time; i.e., $\dot{\varphi}_2 \geqq \dot{\varphi}_1$. Otherwise, core 2 will not have completed switching by the time core 1 is fully switched, and, therefore, the transfer will be incomplete. This requirement can be expressed as

$$i_a \geqq I_A - ni_a . \tag{1}$$

Equation (1) expresses the fact that the resultant mmf acting on core 2 must equal or exceed that acting upon core 1. The assumption is made that $i_a$ remains constant during switching, this following from the model — i.e., $\dot{\varphi}$ = a constant — when a constant-current drive, $I_A$, is applied to produce the reversal.

Furthermore, considering emf's around the winding

$$i_a = \frac{n(\dot{\varphi}_1)_a - (\dot{\varphi}_2)_a}{R},$$

where the subscripts 1 and 2 refer to cores 1 and 2 respectively. For the limiting case where $(\dot{\varphi}_2)_a = (\dot{\varphi}_1)_a$,

$$i_a = \frac{(n-1)(\dot{\varphi}_1)_a}{R},$$

or, inserting the expression for $\dot{\varphi}$ given by the model, i.e., $\dot{\varphi} = \Phi/\tau$,

$$i_a = \frac{(n-1)}{R} \frac{4\pi\Phi}{s_w L} (I_A - ni_a - I_0), \tag{2}$$

where $I_0$ is the threshold current for switching between legs 1 and 3.

Thus, combining (1) with (2) to eliminate $i_a$ and assuming that $I_A \gg I_0$, we obtain as a necessary condition

$$R \leqq \frac{4\pi\Phi}{s_w L} (n-1). \tag{3}$$

We must now consider the requirements of the $P$ phase. In this case

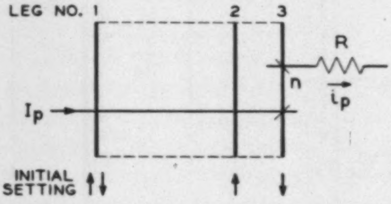

Fig. 25 — Shift register circuit, $P$ phase.

$I_p$, Fig. 25, must be limited to below the threshold for switching between legs 1 and 3:

$$I_p < I_0. \tag{4}$$

The current induced in the coupling winding during switching in this phase is limited only by the resistance of the winding

$$\begin{aligned} i_p &= \frac{n(\dot{\varphi}_1)_p}{R} \\ &= \frac{n}{R}\frac{4\pi\Phi}{s_w l}(I_p - ni_p - i_0), \end{aligned} \tag{5}$$

where $i_0$ is the threshold for switching between legs 2 and 3, and is equal to $I_0 l/L$, $l$ and $L$ being the lengths of switching paths around the small and large holes respectively. Combining (4) and (5), and noting that the reciprocal switching speed $1/\tau_p = 4\pi(I_p - ni_p - i_0)/s_w l$, it follows that

$$\frac{1}{\tau_p} < \frac{I_0\left(1 - \frac{l}{L}\right)}{\frac{s_w l}{4\pi} + \frac{\Phi n^2}{R}},$$

or, taking into account the limitation on $R$, in (3), the limiting speed obtained by making $L \gg l$ is given by,

$$\tau_p > \frac{s_w}{H_0}\frac{n^2}{n - 1}, \tag{6}$$

where $H_0 = 4\pi I_0/L$ is the threshold field for the material. Thus, for a large length ratio, $L/l$, increased speed is obtained by decreasing the switching time constant or increasing the threshold field, both charactersitics of the material only. The turns ratio, $n$, should be kept as small as compatible with the gain requirements. The minimum value for $\tau_p$ estimated in this way has the order of magnitude 10 microseconds for available materials in agreement with the measured range. As noted previously, a further increase in switching speed can be obtained by using an enabling winding.

It is of interest to determine, in a similar manner, the minimum switching speed during the slow phase of a related circuit described recently by Prywes.[17] (The reader is referred to the original article for a full description of the operation of this circuit.) For the present purposes, we need only consider that phase of the operation shown in Fig. 26. Core 1 contains a one. When drive $I$ is applied to produce a reversal between legs 3 and 4 in core 1, the induced current $i$ produces a reversal between legs 1 and 3 in core 2, advancing the one.

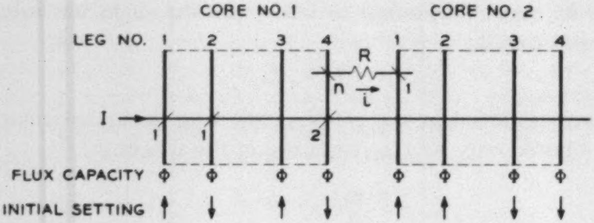

Fig. 26 — One phase of operation of circuit described by Prywes.[17]

As noted previously, it is necessary that $\dot{\varphi}_2 \geqq \dot{\varphi}_1$. However, in this case the fastest transfer of flux to core 2 is obtained by letting $R \to 0$. In this case,

$$\dot{\varphi}_2 \to n\dot{\varphi}_1 \qquad \text{and} \qquad i \to \frac{2nI + I_0}{1 + n^2}.$$

Therefore,

$$\begin{aligned}\frac{1}{\tau^2} &= \frac{4\pi}{s_w L}(i - I_0)\\ &< \frac{4\pi}{s_w L}\left(\frac{2nI + I_0}{1 + n^2} - I_0\right).\end{aligned}$$

In this circuit, in order to prevent an improper conversion of a zero into a one at this stage, it is necessary that $I \leqq I_0$. Therefore,

$$\tau_p > \frac{s_w}{H_0}\left(\frac{1 + n^2}{2n - n^2}\right).$$

Consequently, this arrangement is subject to similar speed limitations. In addition, in this case it is necessary that $n < 2$, thereby limiting the practical possibilities for fanout.

There are a number of alternative ways to organize these circuits. For example, the output of an information core containing a one can be used to inhibit rather than produce switching. However, the alternative arrangements are subject to essentially the same speed limitations.

At the beginning of Section IV it was pointed out that a combinational function could be generated at a given stage of the advance sequence by modifying the input structure of the core. One possibility is illustrated in mirror image form in Fig. 27. Legs 2 and 3 correspond to those of the same numbers in Fig. 10(e). However, the input leg [leg 1 in Fig. 10(e)] is replaced by a number of variable rungs $V_N$, there being as many legs,

$N$, as there are product terms in the Boolean function in its conjunctive canonical form. Each variable leg has the same flux capacity, $\Phi$, as legs 2 and 3. In addition, a bypass leg $B$ having a flux capacity equal to $(N - 1)\Phi$ must be added, and a closure leg $C$ (capacity $\Phi$) must also be added to complete the required flux path. There is a large aperture between legs $B$ and 2, small apertures between the remainder. The "clear" flux setting of the structure is as shown in Fig. 27. Information is advanced into this core by $N$ input pulses, each provided by the output of a preceding information core during an advance phase $A_1$. Each pulse switches flux down in each of the variable rungs. If all $N$ inputs are present simultaneously, a full flux reversal will be produced in both legs $B$ and 2. Each input pulse should be of duration longer than the longest switching time in order to prevent back-switching. The core can be "primed" and then "cleared" by $I_{A_2}$ to advance a one out of the core as in the circuit of Fig. 10(e). On the other hand, if one or more of the inputs are missing, leg $B$ will absorb all flux reversals and rung 2 will remain set downwards. In this case, the core cannot be "primed", and so a zero will be advanced during the $A_2$ phase. It will be clear that the input structure performs an AND function. Any OR terms necessary can be obtained by combining the outputs of separate cores. Thus, to the gain feature provided by the transfluxor output there is added provision for generating any combinational function. The same function can be generated by combining a number of the regular two-aperture cores, but a modified structure may offer a more convenient arrangement. Modified structures of this nature have been included in experimental circuitry, and have been shown to perform satisfactorily.

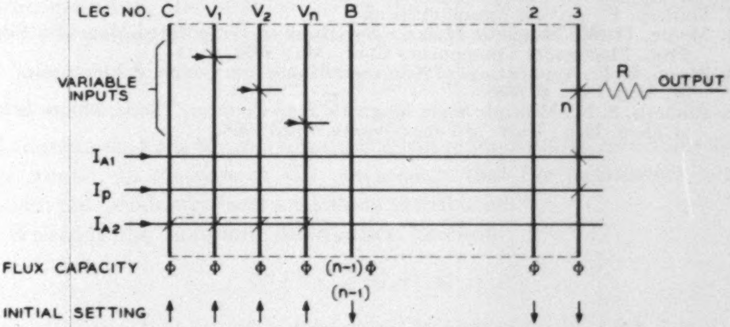

Fig. 27 — Structure for generation of combinational function at a given stage of the advance sequence by modifying input structure of core.

REFERENCES

1. Wang, A. and Woo, W. L., Static Magnetic Storage and Delay Line, J. Appl. Phys., **21,** 1950, p. 49.
2. Karnaugh, M., Pulse Switching Circuits Using Magnetic Cores, Proc. I.R.E., **43,** 1955, p. 570.
3. Guterman, S., Kodis, R. D. and Ruhman, S., Logical and Control Functions Performed with Magnetic Cores, Proc. I.R.E., **43,** 1955, p. 291.
4. Minnick, R. C., Magnetic Switching Circuits, J. Appl. Phys., **25,** 1954, p. 479.
5. Rajchman, J. A. and Crane, H. D., Current Steering in Magnetic Circuits, Trans. I.R.E., **EC6,** 1957, p. 21.
6. Rajchman, J. A., Magnetic Switching, W.J.C.C. Proc., March 1958, p. 63.
7. Hegyi, I. J., Ferromagnetic Spinels with Rectangular Hysteresis Loops, J. Appl. Phys., **25,** 1954, p. 176.
8. Menyuk, N. and Goodenough, J. B., Magnetic Materials for Digital Computer Components, J. Appl. Phys., **26,** 1955, p. 8.
9. Rajchman, J. A. and Lo, A. W., The Transfluxor, Proc. I.R.E., **44,** 1956, p. 321.
10. Gianola, U. F. and Crowley, T. H., The Laddic — A Magnetic Device for Performing Logic, B.S.T.J., **38,** 1959, p. 45.
11. Abbott, H. W. and Suran, J. J., Multihole Ferrite Core Configurations and Applications, Proc. I.R.E., **45,** 1957, p. 1081.
12. Lockhart, N. F., Logic by Ordered Flux Changes in Multipath Cores, I.R.E. Nat. Conv. Rec., **6,** Part 4, 1958, p. 268.
13. Stabler, E. P., Square Loop Magnetic Logic Circuits, General Electric Tech. Info. Series, No. R59ELS8, 1959.
14. Bozorth, R. M., *Ferromagnetism*, D. Van Nostrand Co., New York, 1956, p. 495.
15. Russell, L. B., Diodeless Magnetic Core Logical Circuits, I.R.E. Nat. Conv. Rec., **5,** Part 4, 1957, p. 106.
16. Crane, H. D., A High-Speed Magnetic Logic System Using Magnetic Elements and Connecting Wire Only, Proc. I.R.E., **47,** 1959, p. 63.
17. Prywes, N. S., Diodeless Magnetic Shift Registers Utilizing Transfluxors, Trans. I.R.E., **EC7,** 1958, p. 316.
18. Goto, E., The Parametron, A Digital Computing Element Which Utilizes Parametric Oscillation, Proc. I.R.E., **47,** 1959, p. 1304.
19. Proebster, W. E., Ferroresonant Switching Circuits, Proc. Conf. Nonlinear Magnetics and Magnetic Amplifiers, A.I.E.E., August 1958, p. 434.
20. Hofmockel, D. L., private communication.
21. Donohoe, D. C. and Schulte, H. J., private communication.
22. Bobeck, A. H. and Fischer, R. F., Reversible Diodeless Twistor Shift Register, J. Appl. Phys., **30,** 1959, p. 435.
23. Mallery, P., private communication.
24. Moore, D. W., Magnetic Domain Switching in Evaporated Magnetic Films, Proc. Electronics Components Conf., May 1959, p. 11.
25. Storm, H. F., Applications of Nonlinear Magnetics, Comm. & Electronics, No. 37, May 1958, p. 386.
26. Einhorn, S. N., Multiple State Magnetic Step Counter, Thesis, Moore School of Elect. Eng., Univ. of Pennsylvania, April 1958.

# Amplitude Distribution of Shot Noise

By E. N. GILBERT and H. O. POLLAK

(Manuscript received November 9, 1959)

*A shot noise, $I(t)$, is a superposition of impulses occurring at random Poisson distributed times $\cdots, t_{-1}, t_0, t_1, t_2, \cdots$. In the simplest case, if the impulses all have the same shape $F(t)$, then $I(t) = \sum_i F(t - t_i)$. We study, in this and more general cases, the distribution function $Q(I) = Pr[I(t) \leqq I]$. One of our results is an integral equation for $Q(I)$. This yields explicit expressions for $Q(I)$ in a number of cases, including $F(t) = e^{-t}$; it also permits a computational technique which is applied to $F(t) = e^{-t} \sin \omega t$ for $\omega \gg 1$.*

## I. INTRODUCTION

A shot noise, $I(t)$, is a superposition of impulses occurring at random times $\cdots, t_{-1}, t_0, t_1, t_2, \cdots$. If the impulses all have the same shape, $F(t)$, then

$$I(t) = \sum_i F(t - t_i). \tag{1}$$

More generally, the impulse shapes may be randomly chosen from a family of shapes, $F(a, t)$, depending on a parameter $a$. Then

$$I(t) = \sum_i F(a_i, t - t_i). \tag{2}$$

We assume that the times $t_i$ form a Poisson sequence with rate $n$ impulses per second. In the case of (2), we assume that the parameters $a_i$ are chosen independently from a common distribution.

We study the amplitude distribution function

$$Q(I) = \Pr[I(t) \leqq I].$$

Rice[1] (Section 1.4) considered the noise (1) and noises (2) with $F(a, t) = aF(t)$. He expressed the density function $P(I) = Q'(I)$ as a Fourier integral. The Fourier integral is difficult to evaluate except by means of a series that Rice derived for the case of noises that are nearly gaussian (large impulse rate $n$). In our treatment we derive for (1) an integral equation

$$\int_{-\infty}^{I} x\, dQ(x) = n \int_{-\infty}^{\infty} Q[I - F(t)]F(t)\, dt \tag{3}$$

or, equivalently,

$$IQ(I) = \int_{-\infty}^{I} Q(x)\, dx + n \int_{-\infty}^{\infty} Q[I - F(t)]F(t)\, dt. \tag{4}$$

We solve (3) for some special choices of $F(t)$ illustrated in Fig. 1. The section numbers on Fig. 1 give the part of the text in which each $F(t)$

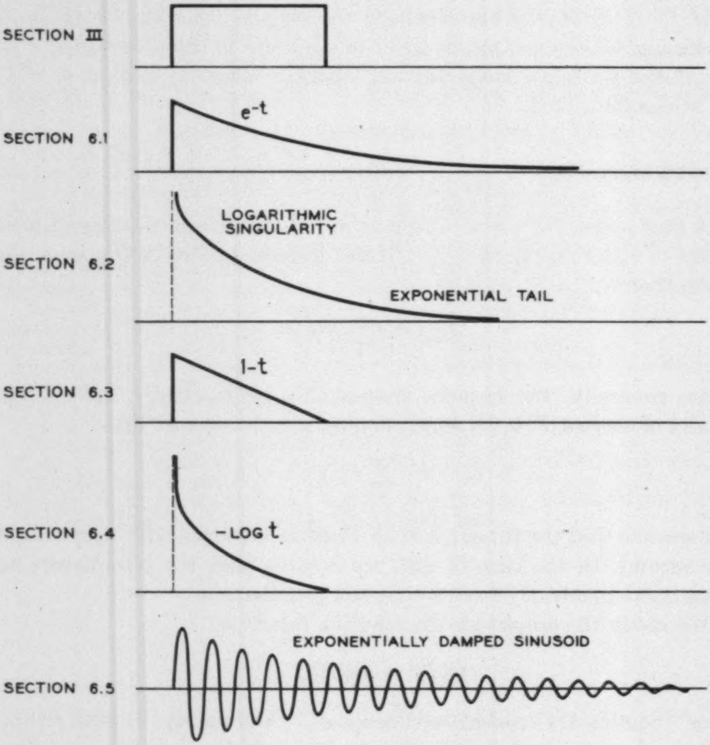
SECTION III
SECTION 6.1
e^-t
LOGARITHMIC SINGULARITY
SECTION 6.2
EXPONENTIAL TAIL
1-t
SECTION 6.3
SECTION 6.4
-LOG t
EXPONENTIALLY DAMPED SINUSOID
SECTION 6.5

Fig. 1 — Impulse functions $F(t)$.

is considered. Analytic solutions are obtained for all cases except the last one of these; this case is an important one in practice, and so was chosen to illustrate a numerical solution of (3).

For purposes of finding $Q(I)$ any given noise of form (2) can be replaced by an equivalent one of form (1). This equivalence is discussed in Section III.

Several different ways of deriving (3) are possible. We give an analytic proof (Section IV) and a probabilistic proof (Section V).

## II. CHARACTERISTIC FUNCTIONS

All impulse functions $F(t)$, $F(a, t)$ will be assumed integrable over $-\infty < t < \infty$. This assumption is no practical restriction and is made to ensure that the series (1) and (2) converge.

We begin by deriving the characteristic function $C(s)$ of $I(t)$ for the noise (2); i.e.,

$$C(s) = E[e^{-sI(t)}]. \tag{5}$$

Here $E$ denotes the expected value.

The characteristic function $C_0(s)$ in the case of the noise (1) is

$$C_0(s) = \exp\left\{-n\int_{-\infty}^{\infty}(1 - \exp[-sF(t)])\,dt\right\} \tag{6}$$

(see Ref. 1, Section 1.4–7). We may obtain $C(s)$ directly from (6) by regarding the noise (2) as a superposition of noises of the form (1) with different choices of $F(t)$. Suppose, for example, that the parameter $a$ has only a discrete range of possible values ($a = A_1, A_2, A_3, \cdots$), and let $p_k$ be the probability of picking $a$ to be $A_k$. Then, collecting together the terms of (2) for which $a$, has the same value, one expresses $I(t)$ as a sum of new independent random variables

$$I(t) = I_1(t) + I_2(t) + \cdots,$$

where $I_k(t)$ is a noise of the form (1) in which the impulse shape is $F(t) = F(A_k, t)$ and the impulses arrive at an average rate, $np_k$ per second. If $C_k(s)$ is the characteristic function of $I_k(t)$,

$$C(s) = E[e^{-s(I_1(t)+I_2(t)+\cdots)}],$$

$$C(s) = C_1(s)C_2(s)\cdots. \tag{7}$$

In (7) each $C_k(s)$ may be evaluated by an expression of the form (6), and the final result is

$$C(s) = \exp\left[-n\left(\int_{-\infty}^{\infty}\{1 - E[e^{-sF(a,t)}]\}\,dt\right)\right]. \tag{8}$$

In (8) the expectation $E$ is taken with respect to the random parameter $a$. Although our derivation of (8) used the assumption that $a$ had a discrete range of values, a convincing limiting argument can be given for the truth of (8) in general. Alternatively one can rederive (8) in general by a slight modification of Rice's derivation of (6).

In a similar way, we find

$$E\,\{\exp[-\,s_1 I(\tau_1) - \cdots - s_N I(\tau_N)]\}$$
$$= \exp\left\{-n \int_{-\infty}^{\infty} 1 - E\left[\exp - \sum_{k=1}^{N} s_k F(a, t - \tau_k)\, dt\right]\right\},$$

which might be used to study the joint distribution of $I(\tau_1), \cdots, I(\tau_N)$.

### III. EQUIVALENCE

In (6) it is evident that there are many different ways of choosing an $F(t)$ to obtain the same distribution of $I(t)$. The integral in (6) remains unchanged if $F(t)$ is replaced by any other function $F_0(t)$ such that, for every choice of $u_1$ and $u_2$, the two sets $S$ and $S_0$ of times $t$ that are defined by:

$$\begin{aligned} S: &\quad u_1 < F(t) \leqq u_2, \\ S_0: &\quad u_1 < F_0(t) \leqq u_2 \end{aligned} \tag{9}$$

have the same measure: For example, the second function in Fig. 1 may be replaced by $e^{-2|t|}$. Some idea of the freedom with which one can construct such a new $F_0(t)$ from a given $F(t)$ may be had from Fig. 2. Given $F(t)$ and $n$, one can construct a measure, $dg(u)$, on the real $u$ line by defining the measure of the interval $u_1 < u \leqq u_2$ to be $n$ times the Lebesgue measure of the set of times $t$ for which $u_1 < F(t) \leqq u$. Then, changing the variable of integration in (6) from $t$ to $u$, one obtains

$$C_0(s) = \exp\left[-\int (1 - e^{-su})\, dg(u)\right]. \tag{10}$$

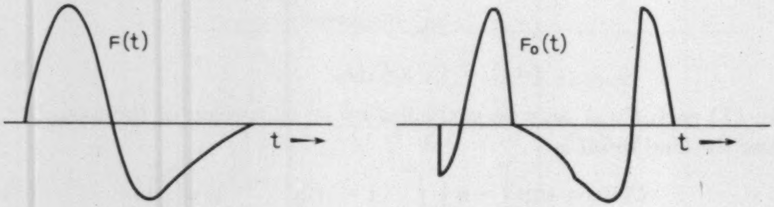

Fig. 2 — A pair of equivalent impulse functions.

Similarly, in the case of noise (2), let $dg(u)$ be defined in such a way that the interval $u_1 < u \leqq u_2$ has measure equal to $n$ times the expected value of the Lebesgue measure of the random set of times $t$ satisfying

$$u_1 < F(a, t) \leqq u_2 . \tag{11}$$

Changing the variable of integration in (8) from $t$ to $u$, one obtains for $C(s)$ an expression that is just the right-hand side of (10). Thus, we call two noises which have the same $dg(u)$ *equivalent* and have shown

*Theorem: The distributions of the amplitudes $I(t)$ of equivalent noises are the same.*

Given a noise (2) with measure $dg(u)$, one can find a function $F(t)$ such that

$$F\left(\int_u^\infty \frac{dg(w)}{n}\right) = u \qquad \text{for } u > 0 \tag{12}$$

and

$$F\left(-\int_{-\infty}^u \frac{dg(w)}{n}\right) = u \qquad \text{for } u < 0. \tag{13}$$

The noise (1) with this choice of $F(t)$ is equivalent to the given noise (2). Then every noise (2) is equivalent to a noise of form (1). For the problem of finding the amplitude distribution we now need consider only noises of the form (1).

As a very simple application of the theorem, consider (2) with the family of impulse functions

$$F(a, t) = \begin{cases} b & \text{if } 0 \leqq t \leqq a \\ 0 & \text{otherwise,} \end{cases} \tag{14}$$

where the parameter $a$ is distributed over positive values only. To find the measure $dg(u)$, note that the set (11) has Lebesgue measure

$$\int_{u_1}^{u_2} dg(u) = \begin{cases} \infty & \text{if } u_1 < 0 \leqq u_2 \\ a & \text{if } 0 \leqq u_1 < b \leqq u_2 \\ 0 & \text{otherwise.} \end{cases}$$

Then, $dg(u)$ must lump all its measure onto two points $u = 0$ and $u = b$. The measure of 0 is $\infty$ and the measure of $b$ is $nA$, where $A = E(a)$, the expected length of the pulse (14). A noise of the form (1) that also has the measure $dg(u)$ is the one with

$$F(t) = \begin{cases} b & \text{if } 0 \leqq t \leqq A \\ 0 & \text{otherwise.} \end{cases} \tag{15}$$

This happens to be a noise for which the amplitude distribution is easily obtained. Since $I(t)$ is just the number of impulses which arrive in the time interval from $t - A$ to $t$, $I(t)$ has the Poisson distribution

$$\Pr\,[I(t) = kb] = \frac{(nA)^k}{k!}\, e^{-nA}.$$

By the theorem, this result also solves the distribution problem for the original noise (14). In the same way, the example may be generalized as follows: If $F(a, t)$ has the form $S(t/a)$ for some given impulse function $S(t)$, then the amplitude $I(t)$ in (2) has the same distribution as the amplitude in (1) where $F(t)$ is taken to be $S[t/(E \mid a \mid)]$.

Although the measure $dg(u)$ determines the amplitude distribution, it does not determine all the statistical properties of the noise. One can easily find examples of functions $F(t)$ and $F_0(t)$ for which the corresponding noises (1), although equivalent, have different joint distributions for the pair of random variables $I(\tau_1)$, $I(\tau_2)$. The spectrum of $I(t)$ is proportional to the squared magnitude of the Fourier transform of $F(t)$ (see Campbell's theorem in Ref. 1, Sections 1.2, 1.3), and so can be changed without changing $dg(u)$.

## IV. DERIVATION

A proof of (3) can be given from the formula (10) for the characteristic function $C(s)$. We will assume here that the impulses $F(t)$ or $F(a, t)$ in question are nonnegative functions. This restriction will be removed in the next section and is made now in order to allow Laplace transform methods to be used.

We have

$$\begin{aligned} C(s) &= E[e^{-sI(t)}] \\ &= \int_{-0}^{\infty} e^{-sI}\, dQ(I) \\ &= s \int_{0}^{\infty} e^{-sI} Q(I)\, dI, \end{aligned}$$

where the last formula is obtained by integrating by parts and noting $Q(0-) = 0$. Thus, $C(s)/s$ is the Laplace transform of $Q(I)$.

Using (10),

$$\begin{aligned} \frac{C'(s)}{C(s)} &= \frac{d}{ds} \log C(s) \\ &= -\int e^{-su} u\, dg(u). \end{aligned}$$

Then

$$-\frac{C'(s)}{s} = \frac{C(s)}{s}\int e^{-su}u\,dg(u). \tag{16}$$

Take the inverse Laplace transform of (16). The product on the right of (16) transforms into a convolution. One of the terms, $C(s)/s$, is the Laplace transform of $Q(I)$. The other term is already in the form of a Laplace transform. Then the product transforms into

$$\int_0^I Q(I-u)u\,dg(u).$$

This integral is another way of writing

$$\int Q[I-F(t)]F(t)\,dt.$$

To prove (3), it now remains to show that the term $-C'(s)/s$ in (16) is the Laplace transform of

$$\int_0^I x\,dQ(x).$$

This follows because

$$C'(s) = -\int_0^\infty e^{-sI}I\,dQ(I)$$

and because $1/s$ is the Laplace transform of the unit step function.

The integral equation (3) can be proved in several other ways; in particular, a more probabilistic proof will be obtained as a byproduct in the next section.

## V. SUMS OF NOISES

Let $I_1(t)$ and $I_2(t)$ be two independent shot noises with impulse responses $F_1(t)$, $F_2(t)$, impulse rates $n_1$, $n_2$ and measures $dg_1(u)$, $dg_2(u)$. Their sum $I(t) = I_1(t) + I_2(t)$ is a shot noise of the form (2). The impulse rate for the sum $I(t)$ is $n = n_1 + n_2$; the function $F(a, t)$ is chosen to be $F_1(t)$ or $F_2(t)$ with probabilities $n_1/n$ and $n_2/n$. The measure $dg(u)$ for $I(t)$ is

$$dg(u) = dg_1(u) + dg_2(u).$$

Since $I(t)$ is the sum of the independent random variables $I_1(t)$ and

$I_2(t)$, the distribution of $I(t)$ can be obtained from those of $I_1(t)$ and $I_2(t)$ by a convolution.

This observation may be used occasionally to compute the distribution function for a given noise $I(t)$. Suppose the measure $dg(u)$ of $I(t)$ can be written as a sum $\Sigma\, dg_i(u)$ of measures of noises $I_i(t)$ for which the distribution functions are known. Then $I(t)$ is equivalent to a sum of independent noises $\Sigma\, I_i(t)$ and its distribution function may be obtained by convoluting the distributions of $I_i(t)$ together. Even when an exact decomposition $dg(u) = \Sigma\, dg_i(u)$ is not known one might approximate $dg(u)$ by a sum to get an approximate $Q(I)$.

A particular instance of a decomposition is the following. Let measures $dg^+(u)$ and $dg^-(u)$ be defined by

$$dg^+(u) = \begin{cases} dg(u) & \text{if } u > 0 \\ 0 & \text{otherwise,} \end{cases}$$

$$dg^-(u) = \begin{cases} dg(u) & \text{if } u \leqq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Then $dg(u) = dg^+(u) + dg^-(u)$. At present, our proof of (3) holds only for positive noises. A similar derivation of (3) for negative noises also holds. In the general case we might consider our noise $I(t)$ to be equivalent to a sum $I^+(t) + I^-(t)$ of independent positive and negative noises, and compute $Q(I)$ as a convolution

$$Q(I) = \int Q^+(I - x)\, dQ^-(x).$$

That (3) holds in general now follows from the next lemma.

*Lemma: Let $dg_1(u)$, $dg_2(u)$, $dg(u)$ be measures such that*

$$dg(u) = dg_1(u) + dg_2(u).$$

*Let $Q_1(I)$, $Q_2(I)$ be solutions of the integral equation corresponding to the measures $dg_1(u)$, $dg_2(u)$. Then for measure $dg(u)$ the convolution*

$$Q(I) = \int Q_1(I - w)\, dQ_2(w)$$

$$= \int Q_2(I - x)\, dQ_1(x)$$

*is a solution of the integral equation (3).*

*Proof:*

$$\begin{aligned}\int_{-\infty}^{I} y\,dQ(y) &= \int_{-\infty}^{I}\int_{w} y\,dQ_1(y - w)\,dQ_2(w)\\ &= \int_{-\infty}^{I}\int_{w} [(y - w) + w]\,dQ_1(y - w)\,dQ_2(w)\\ &= \int_{-\infty}^{\infty}\int_{w} Q_1(I - w - u)u\,dg_1(u)\,dQ_2(w)\\ &\quad + \int_{-\infty}^{\infty}\int_{w} dQ_1(y - w)Q_2(I - w)u\,dg_2(u)\\ &= \int_{-\infty}^{\infty} Q(I - u)u\,dg(u).\end{aligned}$$

Taking $dg_1$ and $dg_2$ as $dg^+$ and $dg^-$, the lemma, together with the result of Section IV, completes the derivation of (3).

A different proof of (3) may now be outlined as follows. Consider first the noise with response function (15). The integral equation (3) holds in this special case. For $dg(u)$ gives measure $nA$ to the point $u = b$ and

$$Q(I) = \sum_{kb \leq I} \frac{(nA)^k}{k!} e^{-nA}.$$

Then

$$\begin{aligned}\int_{-\infty}^{I} x\,dQ(x) &= \sum_{kb \leq I} kb\,\frac{(nA)^k}{k!} e^{-nA}\\ &= nAbQ(I - b)\\ &= \int_{-\infty}^{\infty} Q(I - u)u\,dg(u).\end{aligned}$$

Next, any step function $F(t)$ with a finite number of steps has a measure $dg(u) = \Sigma\, dg_i(u)$, where each $dg_i(u)$ concentrates its measure on a single value of $u$ [a level of $F(t)$]. We have just proved that the integral equation holds for each of the $dg_i(u)$ noises. By the lemma we can conclude that (3) holds for all step function noises. By limiting arguments, one might establish (3) more generally.

## VI. EXAMPLES

For certain choices of $F(t)$ the integral equation can be solved easily. Some special cases of this kind will be examined in this section.

### 6.1 *Example 1*

First consider a noise (1) with

$$F(t) = \begin{cases} e^{-t} & \text{if } t \geqq 0 \\ 0 & \text{if } t < 0. \end{cases}$$

We expect a distribution $Q(I)$ that has a density $P(I) = Q'(I)$ and so write the integral equation in the form

$$IP(I) = n \int_0^1 P\,(I - u)\, du = n \int_{I-1}^{I} P(x)\, dx.$$

Differentiating, we obtain

$$IP'(I) - (n - 1)P(I) = -nP(I - 1).$$

To solve this differential difference equation, note first that, when $0 \leqq I < 1$, $P(I - 1) = 0$. Hence, for $0 \leqq I < 1$,

$$P(I) = cI^{n-1},$$

where $c$ is a constant of integration to be determined. For larger values of $I$, the differential difference equation may be converted to an integral form

$$P(I) = I^{n-1}\left[c - n \int_1^I P(x - 1)x^{-n}\, dx\right].$$

Since the integrand is known for $x < 2$, we can determine $P(I)$ for $I < 2$. Next, this result enables us to integrate further to get $P(I)$ for $I < 3$, etc. Clearly, the analytic form of $P(I)$ changes at $I = 1, 2, 3, \cdots$. For example, when $n = 1$, we have

$$P(I) = c \qquad \text{if } 0 \leqq I \leqq 1,$$

$$P(I) = c(1 - \log I) \qquad \text{if } 1 \leqq I \leqq 2,$$

$$P(I) = c\left[1 - \log I + \int_2^I \frac{\log\,(x - 1)}{x}\, dx\right] \qquad \text{if } 2 \leqq I \leqq 3,$$

etc.

Finally, the constant $c$ must be determined by the condition

$$\int_0^\infty P(I)\, dI = 1.$$

The constant can be determined as follows: The Laplace transform

$\hat{p}(s)$ of $P(I)$ is the characteristic function. By (6),

$$\hat{p}(s) = \exp\left(-n \int_0^s \frac{1 - e^{-y}}{y}\, dy\right).$$

This may be rewritten, with the aid of partial integration, as

$$\hat{p}(s) = \exp\Bigg[-n(1 - e^{-s}) \log s + n\int_0^\infty e^{-y} \log y\, dy - n \int_s^\infty e^{-y} \log y\, dy\Bigg]$$

$$= s^{-n}e^{-n\gamma}\{1 + 0[e^{-s(1-\epsilon)}]\} \qquad \text{for any } \epsilon > 0.$$

Thus, for $0 < I < 1$,

$$P(I) = \frac{e^{-n\gamma}}{\Gamma(n)}\, I^{n-1},$$

where $\gamma = 0.577215665 \cdots$ is Euler's constant. Hence

$$c = \frac{e^{-n\gamma}}{\Gamma(n)}.$$

### 6.2 *Example 2*

Our next example concerns a noise (1) with $F(t)$ defined as follows:

$$F(t) = 0 \qquad \text{for } t \leqq 0,$$

and

$$t = \int_{F(t)}^\infty \frac{e^{-y}\, dy}{y} \qquad \text{for } t > 0.$$

This somewhat artificial noise interests us because it has a very simple $P(I)$. A sketch of $F(t)$ is shown in Fig. 1, marked "Section 6.2". For small $t$, $F(t)$ grows large but only logarithmically:

$$F(t) \sim -\log t.$$

For large $t$, $F(t)$ has an exponential tail,

$$F(t) \sim \gamma^t,$$

where $\gamma$ is Euler's constant.

For this noise, $dg(u) = ne^{-u}\, du$ and the integral equation can be put in the form

$$IP(I)e^I = n \int_0^I P(x)e^x\, dx.$$

Differentiating, we obtain a very simple differential equation and the solution is

$$P(I) = \frac{e^{-I} I^{n-1}}{\Gamma(n)} .$$

This solution exhibits a rapid approach of $P(I)$ to gaussian form (with mean $n$ and variance $n$) as $n \to \infty$.

### 6.3 *Example 3*

If we consider the case

$$F(t) = \begin{cases} 1 - t & 0 \leqq t \leqq 1 \\ 0 & \text{otherwise,} \end{cases}$$

then the distribution $Q(I)$ has a jump of $e^{-n}$ at $I = 0$, since $e^{-n}$ is the probability that no point of the Poisson process falls into an interval of unit length. We therefore seek a density function $P(I)$ such that

$$Q(I) = e^{-n} + \int_0^I P(x)\, dx \qquad \text{for } I \geqq 0.$$

The integral equation then becomes

$$IP(I) = \begin{cases} n \int_0^I P(I - y) y \, dy + nIe^{-n} & \text{if } 0 < I \leqq 1 \\ n \int_0^1 P(I - y) y \, dy & \text{if } 1 < I. \end{cases}$$

If $P(I) = R''(I)$, this becomes

$$IR''(I) - nR(I) = \begin{cases} 0 & \text{if } I < 1 \\ -n[R'(I - 1) + R(I - 1)] & \text{if } I \geqq 1. \end{cases}$$

This can be solved recursively. In the first interval $(I < I)$

$$P(I) = \frac{nR(I)}{I} = \sqrt{n}\, e^{-n} \frac{I_1(2\sqrt{nI})}{\sqrt{I}} ,$$

where $I_1$ is the Bessel function, and where the coefficient of $I_1$ has been determined by substitution in the integral equation.

In the general case, if $I > 1$,

$$P(I) = e^{-n} \sum_{k=0}^{[I]} \frac{(-1)^k}{k!} n^{(k+1)/2} (I - k)^{(k+1)/2} I_{k-1}[2\sqrt{n(I - k)}],$$

where $[I]$ is the largest integer $\leqq I$. For example, if $1 < I \leqq 2$,

$$P(I) = e^{-n}\left[\sqrt{\frac{n}{I}}\, I_1(2\sqrt{nI}) - nI_0(2\sqrt{n(I-1)})\right].$$

(Note that $I_{-1} = I_1$.) These formulas can be derived either from the integral equation, or, with rather more courage but actually less work, from (6).

### 6.4 *Example 4*

Choosing

$$F(t) = \begin{cases} -\log t, & 0 < t \leqq 1 \\ 0 & \text{otherwise} \end{cases}$$

provides another simple case. Again, $\Pr(I = 0) = e^{-n}$, and we seek a density $P(I)$, as in Example 3. For values $I > 0$, $P(I)$ satisfies

$$IP(I) = n\int_0^I P(I-u)ue^{-u}\,du + ne^{-n}Ie^{-I}.$$

Now, letting $R(I) = Ie^I P(I)$, we obtain again the differential equation $IR''(I) + nR(I) = 0$, this time for all $I > 0$. The solution is

$$P(I) = e^{-(I+n)}\sqrt{\frac{n}{I}}\, I_1(2\sqrt{nI}) \qquad \text{for } I > 0.$$

### 6.5 *Example 5*

In this example we let $F(t)$ be the response of a simple tuned circuit to an impulse; i.e.

$$F(t) = e^{-t}\sin\left(\frac{2\pi t}{h}\right). \tag{17}$$

Although the period $h$ appears in $F(t)$ as a parameter, the corresponding measure $dg_h(u)$ tends to a limiting measure $dg(u)$ as $h \to 0$. We will solve the integral equation numerically in this limiting case only. We then expect this result to be applicable as a good approximation whenever the tuned circuit has high $Q$; i.e., $h \ll 1$.

To get the limiting measure $dg(u)$ let us examine $F(t)$ in a small neighborhood of a time $t = T$ at which the sinusoid is at a maximum. During the period from $T$ to $T + h$ the exponential $e^{-t}$ changes only by a factor $e^{-h} = 1 + O(h)$. Thus, aside from terms of order $O(h)$, $F(t)$ is

$$F(t) = e^{-T} \cos\left[\frac{2\pi(t - T)}{h}\right]$$

in this period. Given a level $u > 0$, $F(t) \geqq u$ for a time

$$\frac{h}{\pi} \operatorname{arc\,cos}(ue^{T})[1 + O(h)]$$

during the period $T \leqq t \leqq T + h$. For small $h$, we conclude that $F(t)$ lies above level $u$ for a total amount of time approaching

$$\begin{aligned} \int_0^\infty \frac{dg(w)}{n} &= \frac{1}{\pi} \int_0^{-\log u} \operatorname{arc\,cos}(ue^{T})\, dT \\ &= \frac{1}{\pi} \int_u^1 \frac{\operatorname{arc\,cos} Z}{Z}\, dZ \end{aligned}$$

as $h \to 0$. Similarly, setting the amplitude level $u$ at a negative value, we obtain

$$\int_{-\infty}^{u} \frac{dg(w)}{n} = \frac{1}{\pi} \int_{|u|}^{1} \frac{\operatorname{arc\,cos} Z}{Z}\, dZ.$$

The measure $dg(u)$ is now known. We wish to solve the integral equation in which

$$dg(u) = \begin{cases} \dfrac{n \operatorname{arc\,cos} |u|\, du}{\pi |u|} & \text{if } |u| < 1 \\ 0 & \text{if } |u| \geqq 1. \end{cases} \tag{18}$$

In this case, we had to resort to numerical methods. The integral equation might be approximated directly by a system of linear algebraic equations. However, such an approximation would be troublesome in our case because the integral equation is homogeneous. Unless we could guarantee that the approximating system would have determinant zero, there would be no nontrivial solution at all. The procedure that follows avoids this difficulty.

Let measures $dg^{+}(u)$, $dg^{-}(u)$ be defined as in Section V. We will solve the two integral equations with measures $dg^{+}(u)$ and $dg^{-}(u)$ and convolute the two solutions $P^{+}(I)$, $P^{-}(I)$ together to get the desired density $P(I)$. This approach has the advantage that the integral equation expresses $P^{+}(I)$ in terms of only values of $P^{+}$ for arguments $<I$. Thus, one can compute $P^{+}(I)$ approximately by a simple recurrence.

One might start the recurrence computation by assuming a value of $P^+(I)$ for a small $I$; afterward the solution could be normalized to make

$$\int_0^\infty P^+(I)\,dI = 1.$$

However, we can obtain in the Appendix an asymptotic formula for $P^+(I)$ for $I$ near 0, and thereby start the recurrence with nearly correct values of $P^+(I)$.

Figs. 3 and 4 give the results of the computation for this limiting case of an infinitely rapidly oscillating tuned-circuit response. Computations were made for rates of 2.5, 5 and 10 impulses per unit time, where the time scale is determined by the exponential in (17). Fig. 3 shows $P(I)$ for these three cases; Fig. 4 plots $Q(I)$ on log normal paper and compares the result with the gaussian of the same mean and variance. As expected from Rice's theory, the noticeable differences are in the tails of the distributions; in the shot noise, very large values of $I$ are *more* likely than in the corresponding limiting gaussian.

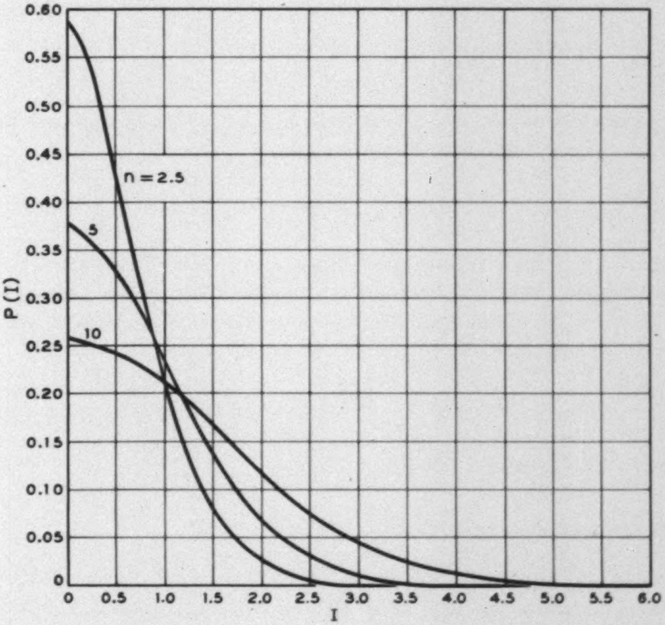

Fig. 3 — Amplitude density $P(I)$ for high-frequency damped sinusoid noise; $F(t) = e^{-t} \sin \omega t$, $\omega \gg 1$; $P(-I) = P(I)$.

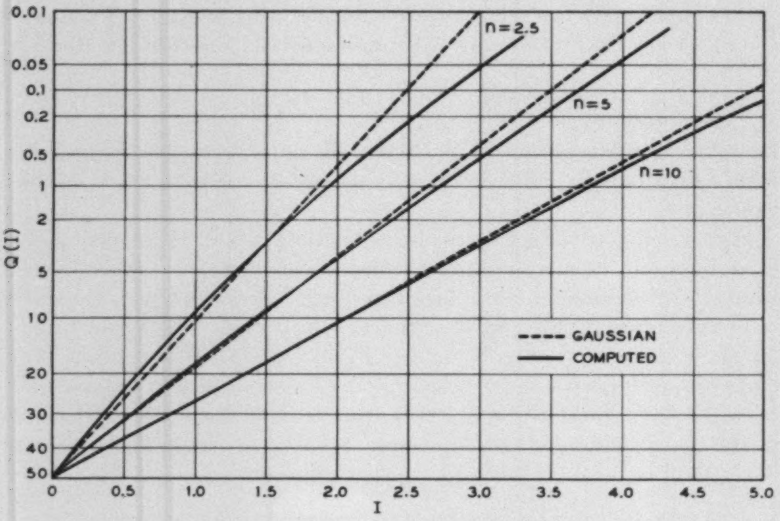

Fig. 4 — Comparison of amplitude distribution $Q(I)$ with gaussian approximation; $Q(-I) = 1 - Q(I)$.

As indicated above, the computation was performed in two stages. First, the integral equation for $P^+(I)$,

$$IP^+(I) = \int_0^I P^+(I - u)\, dg(u),$$

was solved by approximating the integral through Simpson's formula and solving the resulting triangular system of equations; the series for $P^+(I)$ near $I = 0$ was used to start the computation. Then, with the aid of the theory in Section V, $P(I)$ was computed as

$$P(I) = \int_0^\infty P^+(y)P^+(I + y)\, dy,$$

and $Q(I)$ as the integral of $P$.

## VII. ACKNOWLEDGMENT

We wish, at this point, to thank Miss M. A. Lounsberry, for writing a general program for the IBM 704 for solving integral equations of this kind. Without her ingenuity and perseverance the present numerical results would not have been obtained.

## APPENDIX

### *Derivation of the Form for $P^+(I)$ for Small $I$ for a Tuned-Circuit Response*

We begin with the integral equation

$$IP^+(I) = \frac{n}{\pi} \int_0^{\min(1,I)} P^+(I-u) \arccos u\,du.$$

Taking Laplace transforms, we obtain

$$\hat{p}'(s) + \frac{n}{2}\frac{\hat{p}(s)}{s}[1 - I_0(s) + L_0(s)] = 0,$$

where $I_0$ is the modified Bessel function, and $L_0$ the Struve function. Remembering the condition $\hat{p}(0) = 1$, and that $I_0(0) = 1$, $L_0(0) = 0$, we find

$$\hat{p}(s) = \exp\left[-\frac{n}{2}\int_0^s \frac{1 - I_0(u) + L_0(u)}{u}\,du\right].$$

Now $I_0 - L_0$ approaches zero at infinity rapidly enough for that portion of the integral to converge as $s \to \infty$, but the remaining term diverges. We thus write

$$\begin{aligned}
\hat{p}(s) &= \exp\left[-\frac{n}{2}\int_0^s \frac{L_0(u) - I_0(u) + (u+1)^{-1}}{u}\right] du \\
&\qquad + \frac{n}{2}\int_0^s \left[\frac{(u+1)^{-1}}{u} - \frac{1}{u}\right] du \\
&= (s+1)^{-n/2} \exp\left[-\frac{n}{2}\int_0^\infty \frac{L_0(u) - I_0(u) + (u+1)^{-1}}{u}\right] \\
&\qquad \cdot \exp\left[\frac{n}{2}\int_s^\infty \frac{L_0(u) - I_0(u) + (u+1)^{-1}}{u}\,du\right].
\end{aligned}$$

By Ref. 2, p. 426,

$$\frac{I_0(u) - L_0(u)}{u} = \frac{2}{\pi}\int_0^\infty \frac{J_0(x)}{x^2 + u^2}\,dx.$$

Hence,

$$\frac{I_0(u) - L_0(u) - (u+1)^{-1}}{u} = \frac{2}{\pi}\int_0^\infty \frac{J_0(x) - (1+x^2)^{-1}}{x^2 + u^2}\,dx.$$

We may now integrate with respect to $u$ and interchange integrations.

There results

$$\int_0^\infty \frac{I_0(u) - L_0(u) - (u+1)^{-1}}{u}\,du = \int_0^\infty \frac{J_0(x) - (1+x^2)^{-1}}{x}\,dx$$
$$= \int_0^\infty \left[J_1(x) - \frac{2x}{(1+x^2)^2}\right]\log x\,dx$$
$$= \log 2 - \gamma$$

by Gröbner and Hofreiter[3] and Bierens de Haan.[4] Thus,

$$p(s) = e^{-(n/2)\gamma}(2)^{n/2}(s+1)^{-n/2} \cdot \exp\left[\frac{n}{2}\int_s^\infty \frac{L_0(u) - I_0(u) + (u+1)^{-1}}{u}\,du\right].$$

It remains only to discuss the behavior of the last integral for large $s$. By Ref. 2, p. 332,

$$L_0(u) - I_0(u) = -\frac{2}{\pi}\int_0^\infty \frac{\sin ux}{\sqrt{1+x^2}}\,dx.$$

Hence,

$$\int_s^\infty \frac{L_0(u) - I_0(u)}{u}\,du = -\frac{2}{\pi}\int_0^\infty \frac{dx}{\sqrt{1+x^2}}\int_{sx}^\infty \frac{\sin y}{y}\,dy$$
$$= \frac{2}{\pi}\int_0^\infty \log(x + \sqrt{1+x^2})\,\frac{\sin sx}{x}\,dx,$$

while

$$\int_s^\infty \frac{du}{u(u+1)} = \ln\left(1 + \frac{1}{s}\right).$$

Both of these are $O(1/s)$ for large $s$, and hence, for large $s$,

$$\dot{p}(s) = e^{-n/2\gamma}2^{n/2}s^{-n/2} + O[s^{-(n/2)-1}].$$

Hence

$$P^+(I) = \frac{2e^{-(n/2)\gamma}(2I)^{(n/2)-1}}{\Gamma\left(\frac{n}{2}\right)} + O(I^{n/2})$$

for small values of $I$.

## REFERENCES

1. Rice, S. O., Mathematical Analysis of Random Noise, B.S.T.J., **23**, 1944, p. 282.
2. Watson, G. N., *A Treatise on the Theory of Bessel Functions*, Cambridge Univ. Press, Cambridge, 1944.
3. Gröbner, W. and Hofreiter, N., *Integraltafel*, Vol. 2, Springer, Vienna, 1957, p. 531, No. 9.
4. Bierens de Haan, D., *Nouvelles tables d'integrales definies*, Stechert, New York, 1939, p. 140, No. 7.

# On the Recovery of a Band-Limited Signal, After Instantaneous Companding and Subsequent Band Limiting

By H. J. LANDAU

(Manuscript received November 24, 1959)

*If $f(t)$ is a band-limited function, with band limit $-\Omega$ to $\Omega$, the result of instantaneously companding $f(t)$ is in general no longer band-limited. Nevertheless, it has been proved that knowledge of merely those frequencies of the compandor output which lie in the band from $-\Omega$ to $\Omega$ is sufficient to recover the original signal $f(t)$. An iteration formula has been proposed that, in theory, performs the desired recovery. In this paper we study in detail some of the practical questions raised by that formula. We show that the successive approximations converge to the solution $f(t)$ at a geometric rate, uniformly for all $t$, and that the iteration procedure is stable. We then describe a method of performing the recovery in real time and a successful simulation of it on a general-purpose analog computer. The circuit used in the simulation serves as a first approximation to a practical realization of the recovery scheme.*

## I. INTRODUCTION

When a signal, $f(t)$, is transmitted over a channel there is a tendency for the low-amplitude part of $f(t)$ to become masked by the presence of channel noise and for the high-amplitude part of $f(t)$ to become distorted by the nonlinearity of components in those ranges. It would be valuable, therefore, to find a way of assigning to $f(t)$ another signal from which $f(t)$ could be recovered, but which would have the property that its amplitude lay more nearly in the middle ranges than did that of $f(t)$. This second signal is then transmitted, instead of the original $f(t)$. One relatively simple way of obtaining such a signal is by instantaneous companding: The signal sent is $\varphi[f(t)]$, where $\varphi(x)$ is a monotonic function (to allow recovery of $f(t)$ from $\varphi[f(t)]$), which has a large slope around $x = 0$ so as to magnify signals of low amplitude, and which approaches a constant value for large $x$ so as to cut down on signals of high amplitude.

The drawback of instantaneous companding is that it destroys the property of band-limitedness: if $f(t)$ is a band-limited signal with band limit $-\Omega$ to $\Omega$, the signal $\varphi[f(t)]$ is not in general so band-limited. Thus, if the signals are being sent over an idealized band-limiting channel, the function $\varphi[f(t)]$ is distorted in the process of transmission, even though the original $f(t)$ would not have been. This would have put a serious theoretical obstacle in the path of instantaneous companding, were it not for a theorem by Beurling,[1] which shows that, to recover a band-limited function $f(t)$ with band limit $-\Omega$ to $\Omega$ from the companded function $\varphi[f(t)]$, it is not necessary to know the complete spectrum of $\varphi[f(t)]$ but only that part of its spectrum that lies in the frequency interval $-\Omega$ to $\Omega$. More precisely, Beurling has shown that, if $f_1(t)$ and $f_2(t)$ are two band-limited signals with band limit $-\Omega$ to $\Omega$, and if the spectra of $\varphi[f_1(t)]$ and $\varphi[f_2(t)]$ agree in the interval $-\Omega$ to $\Omega$ only, then $f_1(t)$ must coincide when $f_2(t)$. This may be interpreted as saying that "no information is lost" in transmitting $\varphi[f(t)]$ over an idealized band-limiting channel since the result, although bearing no simple relation to $\varphi[f(t)]$, is still sufficient to determine $f(t)$ uniquely. Beurling's proof, however, is nonconstructive, and gives no indication of how the band-limited function $f(t)$ might be obtained from a knowledge of only the part of the spectrum of $\varphi[f(t)]$ between $-\Omega$ and $\Omega$.

In another paper,[2] an iteration formula has been given, by means of which, in theory, the recovery could be performed under the hypothesis (somewhat more restrictive than Beurling's) that $\varphi'(x)$ is bounded, and bounded away from 0. In this paper we will study in detail some of the practical questions raised by that formula. We will show that the successive approximations converge uniformly for all $t$ to the solution $f(t)$ at a geometric rate, and that the iteration procedure is stable. We will then describe a method of instrumenting the recovery in real time and a successful simulation of it on a general-purpose analog computer. The circuit used in the simulation serves as a first approximation to a practical realization of the present recovery scheme.

## II. MATHEMATICAL FUNDAMENTALS

Throughout the subsequent discussions we will be concerned with functions that are square-integrable; this restriction is imposed so that we may pass freely, by means of the Fourier transform, between the time and frequency domains. For a function $f(t)$ that is square-integrable, that is, one for which

$$\int_{-\infty}^{\infty} |f(t)|^2 \, dt < \infty ,$$

the Fourier transform

$$F(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} \, dt \tag{1}$$

is defined and has an inverse given by

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} \, d\omega. \tag{2}$$

Furthermore,

$$\int_{-\infty}^{\infty} |f(t)|^2 \, dt = \int_{-\infty}^{\infty} |F(\omega)|^2 \, d\omega. \tag{3}$$

We will say that a signal $f(t)$ is *band-limited with band* $-\Omega$ *to* $\Omega$ if its Fourier transform $F(\omega)$ vanishes for $|\omega| > \Omega$. Then, by (2), the band-limited signal has a representation as an integral with finite limits

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\Omega}^{\Omega} F(\omega) e^{i\omega t} \, d\omega. \tag{4}$$

Let us next consider an instantaneous compandor, which we will describe by $\varphi(x)$; that is, if $f(t)$ is the input signal to the compandor, the output signal is $\varphi[f(t)]$. We require that the function $\varphi(x)$ satisfy

$$\varphi(0) = 0 \tag{5}$$

and

$$0 < b < \varphi'(x) < B < \infty \qquad (\text{or } -B < \varphi'(x) < -b < 0) \tag{6}$$

in the range of operation for $x$, and we consider the effect of companding on a band-limited function. If $f(t)$ is band-limited with band $-\Omega$ to $\Omega$, the companded signal $\varphi[f(t)]$ need not be band-limited. Nevertheless, it is proved in Ref. 1 that one can compute the original band-limited signal $f(t)$ from a knowledge of merely those frequencies of $\varphi[f(t)]$ which lie in the band from $-\Omega$ to $\Omega$. In order to describe the method of computation, and to enable us to examine the problem in more detail, we introduce the following notation:

i. Let us denote by $T$ the operation of taking the Fourier transform [that is, $Tf(t)$ is the function heretofore denoted by $F(\omega)$], and denote by $T^{-1}$ the operation of taking the inverse Fourier transform [that is, $T^{-1}F(\omega)$ is the function $f(t)$], and let $\chi(\omega)$ be the function that equals 1 for $|\omega| \leqq \Omega$ and equals 0 for $|\omega| > \Omega$. With this notation, we may describe the operation of a low-pass filter on a function $g(t)$ as simply $T^{-1}\chi Tg$, for the action of the filter may be thought of as decomposing $g(t)$ into its frequency components (performing the operation $T$), preserving without change those frequencies in the band $-\Omega$ to $\Omega$ while eliminating all others [multiplying $Tg$ by $\chi(\omega)$], and lastly recomposing the results back into a function of time (performing the operation $T^{-1}$).

ii. For brevity's sake, let us denote by $B$ the space of all band-limited functions with band limit $-\Omega$ to $\Omega$.

iii. Let $Sf = T^{-1}\chi T\varphi[f]$; that is, the operation $S$ applied to a function $f(t)$ consists of companding it and subsequently band-limiting the compandor output. We should observe that, in forming the function $Sf$, we use only those frequencies of the compandor output $\varphi[f]$ which lie in the band $|\omega| \leqq \Omega$. We should also note that $Sf$ is always a band-limited function.

iv. We need some way of measuring distance between two functions. Since we are dealing with square-integrable functions, we choose as our measure the quantity

$$\| f - g \| = \left[ \int_{-\infty}^{\infty} | f(t) - g(t) |^2 \, dt \right]^{\frac{1}{2}}.$$

We refer to $\| f \|$ as "*the norm of the function f*". This norm has many of the properties of ordinary distance; in particular, the triangle inequality

$$\| f + g \| \leqq \| f \| + \| g \|$$

is valid in it. For general square-integrable functions, convergence in norm need not imply ordinary pointwise convergence; that is, we may have functions $f_n(t)$ for which $\| f_n \| \to 0$ but which themselves do not approach 0 at a point (for example, functions with high but thin spikes). It is very important, however, that, for functions in the space $B$, convergence in norm does imply uniform convergence on the whole $t$-axis; indeed we have, if $f(t)$ is in $B$,

$$| f(t) | \leqq \sqrt{\frac{\Omega}{\pi}} \, \| f \|, \qquad \text{for all } t. \tag{7}$$

The proof of (7) is straightforward and is given in Appendix B.

The theorem proved in Ref. 2 asserts that, if we choose a constant $c$ so that

$$| 1 - c\varphi'(x) | \leqq r < 1 \tag{8}$$

for $x$ in the range of operation of the compandor, then, for any two functions $f_1(t)$ and $f_2(t)$ both in $B$, we have

$$\| cSf_1 - cSf_2 - (f_1 - f_2) \| \leqq r \, \| f_1 - f_2 \|. \tag{9}$$

Inequality (9) has many consequences. One of these, established in Ref. 2, is that, given any function $a(t)$ in $B$, the sequence of functions $g_k(t)$, defined iteratively by

$$g_{k+1}(t) = ca(t) + g_k(t) - cSg_k, \qquad \text{with} \quad g_0(t) = 0, \tag{10}$$

converges uniformly on the whole $t$-axis to a limit $g(t)$, which is also in the space $B$. By taking the limit on both sides of (10) we then obtain

$$g(t) = ca(t) + g(t) - cSg,$$

or

$$Sg = a(t). \tag{11}$$

It follows from (9) as well that $g(t)$ is the only function $B$ for which (11) can hold.

Let us interpret this result in physical terms. We may think of a compandor into which is sent a signal $f(t)$ of $B$, and whose output $\varphi[f(t)]$ (which is not in general band-limited) is transmitted over equipment that acts as a pure band-limiter. We may thus describe the received signal as $a(t) = T^{-1}\chi T\varphi[f] = Sf$, and our objective is to recover from $a(t)$ the original compandor input $f(t)$. The iteration formula (10) applied to $a(t)$ does precisely this, for the functions $g_k(t)$ generated by it converge uniformly to a function $g(t)$ in $B$ for which $Sg = a(t)$, and, since there can be only one such function, $g(t)$ must be precisely the desired $f(t)$. The iteration process itself is interpretable in physical terms: the operation $S$, which has to be applied to $g_k$ in order to compute $g_{k+1}$, consists of companding $g_k$ and sending the resulting signal through a filter whose action duplicates that of the transmission network. In essence, $g_{k+1}$ consists of $g_k$ corrected by an appropriate constant multiple of the difference between the received signal $a(t)$ and the signal that would be received, if $g_k(t)$ were companded and transmitted.

In thinking of applying an iteration scheme, the questions of rapidity of convergence and of stability at once present themselves. Let us next consider these.

## III. RAPIDITY OF CONVERGENCE

We will begin by showing that the approximating functions $g_k(t)$ converge to their limit $g(t)$ at a geometric rate. Since the function $g(t)$ in $B$ to be recovered is given by

$$g(t) = \lim_{n \to \infty} g_n(t),$$

we may write

$$\| g_k(t) - g(t) \| = \| g_k(t) - g_{k+1}(t) + g_{k+1}(t) - g_{k+2}(t) + \cdots \|$$
$$\leqq \| g_k(t) - g_{k+1}(t) \| + \| g_{k+1}(t) - g_{k+2}(t) \| + \cdots . \quad (12)$$

Now, from the definition (10), we have

$$g_{i+1} = ca(t) + g_i(t) - cSg_i ,$$
$$g_i = ca(t) + g_{i-1}(t) - cSg_{i-1} ,$$

whence, by subtraction,

$$\| g_{i+1}(t) - g_i(t) \| = \| cSg_i - cSg_{i-1} - (g_i - g_{i-1}) \|, \quad (13)$$

and, since all the functions $g_i$ are in $B$, the inequality (9) may be applied to the right side of (13) to yield

$$\| g_{i+1} - g_i \| \leqq r \| g_i - g_{i-1} \|, \quad (14)$$

for all $i$. By applying the above relation in turn to $\| g_i - g_{i-1} \|$ and so on down to $\| g_1 - g_0 \| = \| a(t) \|$, we may replace (14) by

$$\| g_{i+1} - g_i \| \leqq r^i \| a \|, \quad (15)$$

which, together with (12), yields

$$\| g_k - g \| \leqq \frac{r^k}{1 - r} \| a \|.$$

Since the function $g_k - g$ is in $B$ we may take advantage of the relation between the norm and absolute value that holds (see point iv in Section II) to conclude that actually

$$| g_k(t) - g(t) | \leqq \frac{\sqrt{\Omega/\pi} \, \| a \|}{1 - r} r^k, \qquad r < 1, \qquad \text{for all } t.$$

This establishes that the convergence of $g_k$ to $g$ is geometric in rapidity over the whole $t$-axis. The constant $r$, which determines the actual convergence rate, comes from (8) and depends only on the companding curve $\varphi(x)$. In order to obtain the fastest convergence, $c$ should be chosen so as to make $r$ as small as possible.

## IV. STABILITY

The stability of an iteration scheme refers to its sensitivity to error. In the case at hand, the solution $g(t)$ is the limit of the functions $g_k(t)$ defined by (10), where we are interpreting $a(t)$ as the signal received when the band-limited signal $g(t)$ is companded and subsequently transmitted over a noiseless band-limiting channel, and where the operation $S$ consists of companding followed by band-limiting. With this model in mind, it is easy to imagine that, in a real application, the received signal would not be, because of noise, precisely $a(t)$; or that the compandor, when applied to $g(t)$, had not acted on the precise curve $\varphi(x)$; or that the channel was not precisely an ideal band-limiting channel. The iteration procedure is *stable* if the function $g^*(t)$ that it produces under each of these three conditions of error differs from the true $g(t)$ by an amount commensurate with the error.

Let us first take up the case that the received function $a^*(t)$ is not equal to $a(t)$. The iteration scheme (10) applied to $a^*(t)$ yields a function $g^*(t)$ in $B$ for which

$$g^* = ca^* + g^* - cSg^*,$$

while the true $g(t)$ in $B$ satisfies

$$g = ca + g - cSg.$$

Subtracting the two equations above and taking the norm of both sides, we obtain

$$\| g^* - g \| \leqq c \| a^* - a \| + \| cSg - cSg^* - (g - g^*) \| .$$

We may now apply (9), obtaining

$$\| g^* - g \| \leqq c \| a^* - a \| + r \| g^* - g \| ,$$

from which

$$\| g^* - g \| \leqq \frac{c}{1 - r} \| a^* - a \|,$$

or, passing to absolute values (as in point iv of Section II),

$$| g^*(t) - g(t) | \leqq \frac{c\sqrt{\Omega/\pi}}{1 - r} \| a^* - a \|, \qquad \text{for all } t. \tag{16}$$

This is precisely a statement of stability, for it asserts that the maximum deviation of $g^*$ from $g$ is bounded by a fixed constant multiple of the norm (in our case the square root of the energy) of the error $a^* - a$.

Let us consider next the effect of a compandor error on the iteration; that is, the possibility that the compandor output is not $\varphi[g(t)]$ but rather $\varphi^*[g(t)]$, where $\varphi^*(x)$ is a curve not identical with $\varphi(x)$. Let us assume that the companding itself is stable; i.e., that the quantity

$$\| \varphi^*[g(t)] - \varphi[g(t)] \| ,$$

which represents the square root of the energy of the difference of the two outputs, is commensurate with the compandor error, which we may measure by the quantity

$$\sup_x \mid \varphi^*(x) - \varphi(x) \mid .$$

What we mean by this precisely is that there exists a fixed constant $K$ such that

$$\| \varphi^*[g] - \varphi[g] \| \leqq K \sup_x \mid \varphi^*(x) - \varphi(x) \mid$$

for all functions $g(t)$ under consideration in the problem. As we have seen in point iii of Section II, the results of transmitting the two outputs are, respectively,

$$a^*(t) = T^{-1}\chi T\varphi^*[g] \qquad \text{and} \qquad a(t) = T^{-1}\chi T\varphi[g]$$

so that, utilizing the stability shown above of the recovery formula with respect to received signals, we have, from (16),

$$\begin{aligned} \mid g^*(t) - g(t) \mid &\leqq \frac{c\sqrt{\Omega/\pi}}{1-r} \| a^* - a \| \\ &= \frac{c\sqrt{\Omega/\pi}}{1-r} \| T^{-1}\chi T\varphi^*[g] - T^{-1}\chi T\varphi[g] \|, \end{aligned} \tag{17}$$

where $g^*(t)$ is the function yielded by the iteration scheme on the basis of the erroneous signal $a^*(t)$. Since, by (3), the Fourier transform of a function has the same norm as the function, we have

$$\| T^{-1}\chi T\varphi^*[g] - T^{-1}\chi T\varphi[g] \| = \| \chi T\varphi^*[g] - \chi T\varphi[g] \|$$

and

$$\| T\varphi^*[g] - T\varphi[g] \| = \| \varphi^*[g] - \varphi[g] \|,$$

while

$$\| \chi T\varphi^*[g] - \chi T\varphi[g] \| \leqq \| T\varphi^*[g] - T\varphi[g] \|,$$

since the two sides of the inequality represent integrals of the same positive function, over a finite and an infinite interval respectively. Combining these with (17), we obtain

$$\mid g^*(t) - g(t) \mid \leqq \frac{c\sqrt{\Omega/\pi}}{1-r} \| \varphi^*[g] - \varphi[g] \|,$$

and, by the assumption of compandor stability, the right-hand side above is commensurate with

$$\sup_x \mid \varphi^*(x) - \varphi(x) \mid .$$

We conclude that the iteration procedure is stable with respect to a compandor error whenever the companding process itself is thus stable.

Lastly we take up the question of stability under a variation of the channel characteristic. That is, we suppose that the compandor output $\varphi[g(t)]$ is transmitted over a channel whose effect on it is $T^{-1}\chi^*T\varphi[g]$ rather than $T^{-1}\chi T\varphi[g]$, where the function $\chi^*(\omega)$ differs from the ideal $\chi(\omega)$ of point i in Section II. We have, from (16),

$$\begin{aligned} | g^*(t) - g(t) | &\leqq \frac{c\sqrt{\Omega/\pi}}{1 - r} \| a^* - a \| \\ &= \frac{c\sqrt{\Omega/\pi}}{1 - r} \| T^{-1}\chi^*T\varphi[g] - T^{-1}\chi T\varphi[g] \| , \end{aligned} \tag{18}$$

and, by (3),

$$\begin{aligned} &\| T^{-1}\chi^*T\varphi[g] - T^{-1}\chi T\varphi[g] \| \\ &\qquad = \| \chi^*T\varphi[g] - \chi T\varphi[g] \| = \| (\chi^* - \chi) T\varphi[g] \| . \end{aligned}$$

Now from the definition of the norm (point iv in Section II), we may estimate the quantity $\| (\chi^* - \chi) T\varphi[g] \|$ in various ways. We may choose to say that

$$\| (\chi^* - \chi) T\varphi[g] \| \leqq \sup_{\omega} | \chi^*(\omega) - \chi(\omega) | \, \| T\varphi[g] \| , \tag{19}$$

from which we can show that, for signals of bounded energy, the recovery computation is stable with respect to a departure of the transmission characteristic from the ideal $\chi(\omega)$, when the deviation is measured by the quantity

$$\sup_{\omega} | \chi^*(\omega) - \chi(\omega) | .$$

That is, we will be able to conclude that the error $| g^*(t) - g(t) |$ will be small if

$$\sup_{\omega} | \chi^*(\omega) - \chi(\omega) |$$

is sufficiently small. The weakness of this sort of stability statement lies in its requirement that $| \chi^*(\omega) - \chi(\omega) |$ be everywhere small; it yields no information when $| \chi^*(\omega) - \chi(\omega) |$ is small everywhere, except on a small segment of the $\omega$-axis. In those cases, the quantity

$$\sup_{\omega} | \chi^*(\omega) - \chi(\omega) |$$

ceases to be an adequate measure of closeness, and we would prefer to have a stability statement involving $\| \chi^* - \chi \|$, for this may be small even when $| \chi^*(\omega) - \chi(\omega) |$ is occasionally large. We may find such a statement, for functions whose frequency spectrum, after companding, is bounded, by using

$$\| (\chi^* - \chi) T\varphi[g] \| \leqq \| \chi^* - \chi \| \sup_{|\omega| < \Omega} | T\varphi[g] | . \tag{20}$$

The weakness in turn of

$$\| \chi^* - \chi \| = \left[ \int_{-\Omega}^{\Omega} | \chi^* - \chi |^2 \, d\omega \right]^{\frac{1}{2}}$$

as a measure of closeness is that it may become large, even when

$$| \chi^* - \chi |$$

is mostly small, simply because the interval of integration $| \omega | \leqq \Omega$ is large. To find an expression for the size of error that has the virtues of both of those above without the disadvantages of either — that is, one which is not sensitive either to occasionally large values of $| \chi^* - \chi |$, or to the length of the band $| \omega | \leqq \Omega$ — we may apply a combination of estimates (19) and (20) to the quantity $\| (\chi^* - \chi) T\varphi [g] \|$. Let us divide the interval $-\Omega$ to $\Omega$ into two complementary sets, $A$ and $A' = [(-\Omega, \Omega) - A]$, and let us define

$$\epsilon_A = \sup_{\omega \varepsilon A} | \chi^* - \chi |,$$

$$\epsilon_A{}' = \left[ \int_{\omega \varepsilon A'} | \chi^* - \chi |^2 \, d\omega \right]^{\frac{1}{2}}.$$

If $| T\varphi[g] |$ for $| \omega | \leqq \Omega$ and $\| \varphi[g] \|$ are bounded by $M$, then, using (19) on the set $A$ and (20) on the set $A'$, we obtain

$$\| (\chi^* - \chi) T\varphi[g] \| \leqq M \min_A \max(\epsilon_A \, , \, \epsilon_A{}').$$

Hence, from (18),

$$| g^*(t) - g(t) | \leqq \frac{c\sqrt{\Omega/\pi}}{1 - r} M \inf_A \max \, (\epsilon_A \, , \epsilon_A{}').$$

This establishes that, for signals whose energy and frequency spectrum after companding, are bounded by a fixed constant, the iteration procedure is stable with respect to a departure of the transmission characteristic from the ideal $\chi(\omega)$; our present measure of deviation is the best combination of $| \chi^*(\omega) - \chi(\omega) |$ and $\| \chi^* - \chi \|$, in the sense of minimizing over all sets $A$ the quantity $\max(\epsilon_A \, , \, \epsilon_A{}')$. The two boundedness restrictions we have imposed do not seem unduly severe.

## V. INSTRUMENTATION OF THE ITERATION FORMULA

With the stability of the recovery computation thus established, we will now describe a way of instrumenting the iteration formula in real time, and a simulation of the resulting recovery process on a general-purpose analog computer. The iteration formula is

$$g_{n+1} = ca(t) + g_n - cSg_n \, , \qquad \text{with} \quad g_0 = 0,$$

and we have already interpreted $a(t)$ as the signal received after the function $g(t)$ in $B$ has been companded and transmitted over a band-limiting channel, and the operation $S$ as companding, followed by band-limiting. The iteration is thus performable by analog methods, with the aid of a compandor and of a band-limiting filter, used to carry out the operation $S$. A filter of this type has a delay, however. Thus, in order to add its output, $Sg_n$, to the function $ca(t) + g_n(t)$, as required by the iteration, the latter would have to be delayed by an amount equal to the filter's delay. We may obviate the necessity for an additional delay network by observing that, since the function $ca(t) + g_n(t)$ is in $B$, passing it through the filter does not distort it, so that the addition may be performed before filtering. Thus, a possible circuit for performing the iteration is that of Fig. 1. By connecting $s$ of these circuits in series, and by supplying as input the function $g_1(t) = a(t)$, the output will be the (delayed) approximation $g_{s+1}(t)$ to $g(t)$, for which

$$|g(t) - g_{s+1}(t)| \leqq \frac{c\sqrt{\Omega/\pi}}{1-r}\|a\|r^{s+1}.$$

The circuit of Fig. 1 served as the basis for a simulation on a general-purpose analog computer. The companding curve $\varphi(x)$ was chosen to be of the type described by Mallinckrodt[3] and consisted of a straight line of slope 10 for $-0.2 \leqq x \leqq 0.2$ that had joined to it at $x = \pm 0.2$ a logarithmic curve which matched it in slope. The range of interest for $x$ was $|x| < 2.5$. The constant $c$ was chosen as 1/12, yielding for $r$ the relatively large value of $r = 14/15$. The band-limiting filter was simulated from an expression kindly supplied by J. Bangert. Since it required 13 integrators, it was not possible to simulate more than one stage of the circuit, so that the iteration was performed step by step; the output $g_{n+1}(t)$ was recorded at every step and served, along with $a(t)$, as the input for the next iteration. Since the simulation was per-

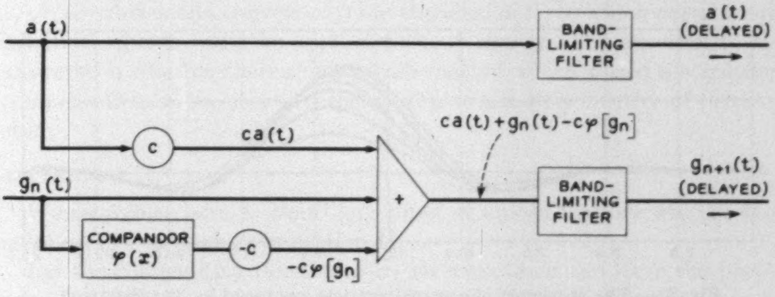

Fig. 1 — Circuit for simulation on a general-purpose analog computer.

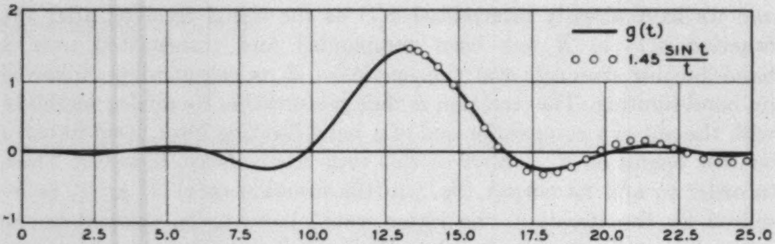

Fig. 2. — The figure $g(t)$ to be recovered.

formed to test the efficacy of the recovery process, the function $g(t)$, to which the approximations $g_n(t)$ were expected to converge geometrically, was picked in advance to be 1.45 sin $t/t$, which is in $B$, as required.

The function actually used for $g(t)$ was an approximation to this, generated on the computer as the step response of the filter, and appears in Fig. 2. Since the filter is not ideal, the $g(t)$ used does not coincide precisely with 1.45 sin $t/t$ but it is a good approximation; the closeness of the two curves provides, further, a measure of the filter's performance. The curve $a(t)$, obtained as a result of companding and filtering the $g(t)$ of Fig. 2, appears as the bottom-most of the curves of Fig. 3. The remaining curves of Fig. 3 represent the odd approximants: $g_3$, $g_5$, $g_7$, $g_9$, $g_{11}$ and $g_{13}$, yielded by the iteration. They are seen to converge well, although their limit is not quite the function $g(t)$. To test whether the error was due simply to the various inherent machine insensitivities, the last approximation, $g_{13}(t)$, was companded and band-limited, and the result was compared with the original $a(t)$. The difference of these two functions appears in Fig. 4 and is seen to be very small. The simulation consequently appears to be successful, in that it verifies, in a special case and within the limits of machine accuracy, the theoretical predictions of convergence and stability for the recovery process.

As we have mentioned before, an obvious way of mechanizing this

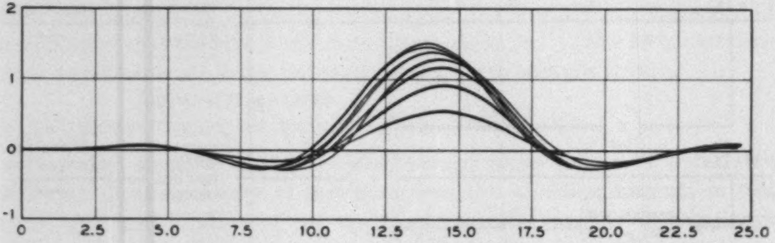

Fig. 3. — The sequence of approximations produced by the iteration.

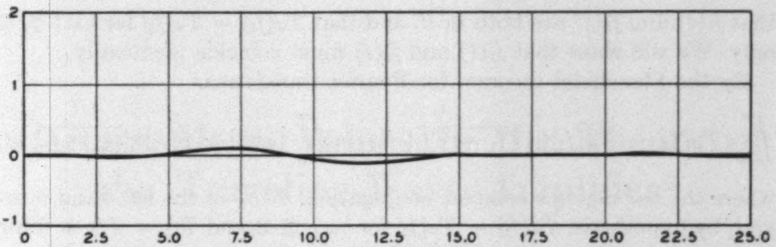

Fig. 4 — Difference between the original $a(t)$ and companded band-limited approximation $g_{13}(t)$.

recovery procedure is to connect $s$ of the circuits of Fig. 1 in series, and to supply as input the function $g_1(t) \equiv a(t)$; the output will then be the approximation $g_{s+1}(t)$ to $g(t)$, but delayed by $s$ times the delay of the band-limiting filter. This delay is an undesirable feature in practice, and may perhaps be decreased, at the expense of some error in the recovery, by using a filter with a smoother cutoff.

## VI. OPEN QUESTIONS

This study has concerned itself until now with the idealized version of the problem — one in which the effect of transmission on the companded signal has been to pass all frequencies $|\omega| \leqq \Omega$ without change and to eliminate all others. We have gone beyond this formulation only to show the stability of the recovery process with respect to a variation of the transmission characteristic from the ideal $\chi(\omega)$; that is, we have shown that, if this variation is not large, the error produced by applying the present recovery procedure will not be large. The problem of how we should proceed when given a signal $a^*(t) = T^{-1}\chi^* T\varphi\,[g]$, with $\chi^*(\omega) = 0$ for $|\omega| > \Omega$ but widely different from $\chi(\omega)$ in the band $|\omega| \leqq \Omega$, remains an open one. We may, of course, precede the recovery by passing $a^*(t)$ through a compensating network with characteristic $1/\chi^*(\omega)$; this would convert $a^*(t)$ to the ideal $a(t)$, to which our present iteration scheme could be applied without change. The question to be answered is whether there exists an alternative, which would not require compensation of the received signal; this is a matter worthy of further study.

## APPENDIX A

We reproduce here A. Beurling's proof of uniqueness; we will use the notation of i through iv in Section II.

Let the companding function $\varphi(x)$ be monotonic and have the property that $\varphi[f(t)]$ is square-integrable whenever $f(t)$ is. Let us also suppose that $f_1(t)$ and $f_2(t)$ are both in $B$, and that $T\varphi[f_1] = T\varphi[f_2]$ for $|\omega| \leqq \Omega$ only. We will show that $f_1(t)$ and $f_2(t)$ must coincide identically.

By the Plancherel theorem for Fourier transforms,

$$\int_{-\infty}^{\infty} \{T\varphi[f_1] - T\varphi[f_2]\}\{\overline{Tf_1 - Tf_2}\}\, d\omega = \int_{-\infty}^{\infty} \{\varphi[f_1] - \varphi[f_2]\}\{\overline{f_1 - f_2}\}\, dt,$$

where the bar denotes complex conjugation. Now in the left-hand integral, by hypothesis, $T\varphi[f_1] = T\varphi[f_2]$ for $|\omega| \leqq \Omega$, and $Tf_1 = Tf_2 \equiv 0$ for $|\omega| > \Omega$, since $f_1$ and $f_2$ are in $B$. Thus,

$$\int_{-\infty}^{\infty} \{\varphi[f_1] - \varphi[f_2]\}\{f_1 - f_2\}\, dt = 0. \tag{21}$$

But, since $\varphi$ is monotonic, the integrand of (21) is nonnegative, for if $f_1(t) \geqq f_2(t)$, then $\varphi[f_1(t)] \geqq \varphi[f_2(t)]$, and, similarly, if $f_1(t) \leqq f_2(t)$, then $\varphi[f_1(t)] \leqq \varphi[f_2(t)]$. Thus (21) implies that $f_1(t) \equiv f_2(t)$.

## APPENDIX B

We will show here that, for functions in the space $B$, convergence in norm implies uniform convergence on the whole $t$-axis.

Let $f_n(t)$ be a sequence of functions in $B$, with $\| f_n(t) \| \to 0$. By applying Schwarz's inequality to the representation (4) we obtain

$$|f_n(t)| \leqq \frac{\sqrt{2\Omega}}{\sqrt{2\pi}} \left[ \int_{-\Omega}^{\Omega} |F_n(\omega)|^2\, d\omega \right]^{\frac{1}{2}},$$

or

$$|f_n(t)| \leqq \sqrt{\Omega/\pi}\, \| F_n(\omega) \|, \tag{22}$$

where $F_n(\omega)$ is the Fourier transform of $f_n(t)$. But, by (3),

$$\| F_n(\omega) \| = \| f_n(t) \|,$$

so that (22) becomes

$$|f_n(t)| \leqq \sqrt{\Omega/\pi}\, \| f_n \|$$

whence we have

$$|f_n(t)| \to 0, \qquad \text{uniformly for all } t.$$

## REFERENCES

1. Beurling, A., private communication (see Appendix A).
2. Landau, H. J., and Miranker, W. L., to be published.
3. Mallinckrodt, C. O., Instantaneous Compandors, B.S.T.J., **30**, July 1951, p. 706.

# Certain Mean Values in the Theory of the Traveling-Wave Amplifier*

By L. A. MacCOLL

(Manuscript received October 26, 1959)

*Some simple results relating to certain mean values are given. It is not assumed that the signals are necessarily small; hence nonlinear effects are taken into account.*

The purpose of this note is to give a few simple results relating to certain mean values that occur in the theory of the traveling-wave amplifier. Whereas all of the previous theory of the amplifier has been based on the assumption that the signals are small, so that the system behaves effectively as a linear system, no such assumption is involved in the results given here.

After some idealization of the physical system, the fundamental equations of the traveling-wave amplifer can be written as follows:[2]

$$L\frac{\partial I}{\partial t} + RI = -\frac{\partial V}{\partial x}, \tag{1}$$

$$C\frac{\partial V}{\partial t} = -\frac{\partial I}{\partial x} + \alpha\frac{\partial \rho}{\partial t}, \tag{2}$$

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho v)}{\partial x} = 0, \tag{3}$$

$$\frac{\partial v}{\partial t} + \frac{1}{2}\frac{\partial v^2}{\partial x} = \beta\frac{\partial V}{\partial x}. \tag{4}$$

The independent variables $t$ and $x$ represent, respectively, time and distance measured in the axial direction from the driving point; $I$, $V$, $\rho$ and $v$ denote the instantaneous local values of the current in the conductor, the potential of the conductor, the linear charge density of the electron stream, and the velocity of the electrons, respectively; $L$, $R$, $C$, $\alpha$ and $\beta$ are constants.

---

* This material was prepared as a memorandum or report in 1946, but was never published. It has, however, been known to people working in the field, and it has been mentioned in published work.[1] It seems desirable that it be made generally available in its original form.

Now suppose that we have a state of the system in which, for each value of $x$, the variables $I$, $V$, $\rho$ and $v$ are all periodic functions of time, with the period $T$. (It is to be noted that nothing is assumed about the waveforms of these periodic functions.)

By (1), we have the relation

$$L \frac{1}{T} \int_{t_0}^{t_0+T} \frac{\partial I}{\partial t}\, dt + R \frac{1}{T} \int_{t_0}^{t_0+T} I\, dt = -\frac{\partial}{\partial x} \frac{1}{T} \int_{t_0}^{t_0+T} V dt, \tag{5}$$

where $t_0$ is an arbitrary constant.

The first term in the left-hand member of (5) vanishes, because

$$\int_{t_0}^{t_0+T} \frac{\partial I}{\partial t}\, dt = I(x, t_0 + T) - I(x, t_0),$$

and because $I$ is periodic with respect to $t$ with the period $T$. The expressions

$$\frac{1}{T} \int_{t_0}^{t_0+T} I\, dt \qquad \text{and} \qquad \frac{1}{T} \int_{t_0}^{t_0+T} V\, dt,$$

which we shall denote by the symbols $\bar{I}$ and $\bar{V}$ respectively, are the means of $I$ and $V$ with respect to $t$ over the period $T$, for an arbitrary value of $x$. A bar over a letter is used in this sense throughout the discussion.

Thus we have the relation

$$R\bar{I} = -\frac{d\bar{V}}{dx}. \tag{6}$$

Similarly, from (2), (3) and (4), we get the relations

$$\frac{d\bar{I}}{dx} = 0, \tag{7}$$

$$\frac{d(\overline{\rho v})}{dx} = 0, \tag{8}$$

$$\frac{d\overline{v^2}}{dx} = 2\beta \frac{d\bar{V}}{dx}. \tag{9}$$

The general solution of (6), (7), (8) and (9) is

$$\begin{aligned} \bar{I} &= K_1, \\ \bar{V} &= -K_1 R x + K_2, \end{aligned}$$

$$\overline{\rho v} = K_3 ,$$

$$\overline{v^2} = -2K_1\beta Rx + 2K_2\beta + K_4 ,$$

where the $K$'s are arbitrary constants.

The most interesting and important state of the system is that in which, at the driving end, we have

$$\bar{I} = 0, \qquad \bar{V} = 0, \qquad \rho = \rho_0 , \qquad v = v_0 , \tag{10}$$

where $\rho_0$ and $v_0$ are constants. In this state, and for any value of $x$, we have the relations

$$\bar{I} = 0, \qquad \bar{V} = 0, \qquad \overline{\rho v} = \rho_0 v_0 , \qquad \overline{v^2} = v_0^2 . \tag{11}$$

This result can be stated in words as follows: If at the driving point the mean values of the conductor current and voltage are zero, and if at the same point $\rho$ and $v$ have the constant values $\rho_0$ and $v_0$ , then the mean values of the conductor current and voltage are everywhere zero, the mean value of the electron convection current is everywhere $\rho_0 v_0$ , and the mean value of the square of the electron velocity is everywhere $v_0^2$.

We note that, although the system is nonlinear, there is no rectification of the applied signals.

By the Schwarz inequality, we have the relation

$$(\overline{\rho v})^2 \leqq (\overline{\rho^2})(\overline{v^2}).$$

This, together with the relations $\overline{\rho v} = \rho_0 v_0$ and $\overline{v^2} = v_0^2$, implies that, in the state to which the equations (10) relate, we have everywhere the relation

$$\overline{\rho^2} \geqq \rho_0^2 . \tag{12}$$

By the Schwarz inequality, we also have the relation

$$(\overline{1 \cdot v})^2 \leqq (\overline{1^2})(\overline{v^2}) = \overline{v^2};$$

and this, together with the relation $\overline{v^2} = v_0^2$, gives us the relation

$$|\bar{v}| \leqq v_0 . \tag{13}$$

REFERENCES

1. Nordsieck, A., Theory of Large-Signal Behavior of Traveling-Wave Amplifiers, Proc. I.R.E., **41,** 1953, p. 630.
2. Pierce, J. R., *Traveling-Wave Tubes*, D. Van Nostrand Co., New York, 1950.

# Radio Frequency Interference Considerations in the TD-2 Radio Relay System

By H. E. CURTIS

(Manuscript received July 23, 1959)

*Relationships are developed between the ratio of a desired RF carrier to an interfering co-channel RF carrier and the telephone channel interference that results therefrom. Objectives are set down in terms of permissible noise in a telephone channel for each individual RF interference on a hypothetical long system. Finally, systems applications of these interference considerations are discussed.*

## I. INTRODUCTION

The number of microwave systems that may be used on a particular frequency assignment in a given area is determined by the mutual interference between the systems. The Bell System is particularly conscious of this problem because of its extensive use of broadband microwave systems.

At the end of 1958 the Bell System had in operation approximately 215,000 one-way broadband channel miles of microwave systems in the United States, corresponding to approximately 7,000 transmitters and an equal number of receivers. A large portion of this extensive network makes use of the TD-2 radio relay system[1,2] operating in the 3700- to 4200-mc band. The TH system[3] operating in the 6-kmc range is under active development, and the TJ system operating in the 11-kmc common carrier band (10,700 to 11,700 mc) is now in commercial use. Equipments made by other manufacturers also operate in the 4- and 6-kmc bands.

In spite of the highly directive antennas now in use, a certain fraction of the transmitter power from any station will be radiated in directions other than that for which it is intended, and this becomes a potential source of intrasystem interference to itself and also to other systems operating in the same frequency band.

This paper is directed specifically toward the TD-2 system because of its extensive use and the experience gained with it. However, the philosophy developed herein may be applied directly to the study of interference arising in or from other systems.

It is sufficient for the purposes of this paper to point out that the TD-2 system, in common with many other microwave systems, uses frequency modulation. It can provide six two-way broadband channels, each of which can handle two one-way television channels or 480 or more two-way telephone channels.

The frequency arrangement at a repeater is shown diagrammatically in Fig. 1. The six channels in one direction operate on channel assignments whose midband frequencies are 80 mc apart. The six channels in the other direction are interlaced with the first six, and thus there are

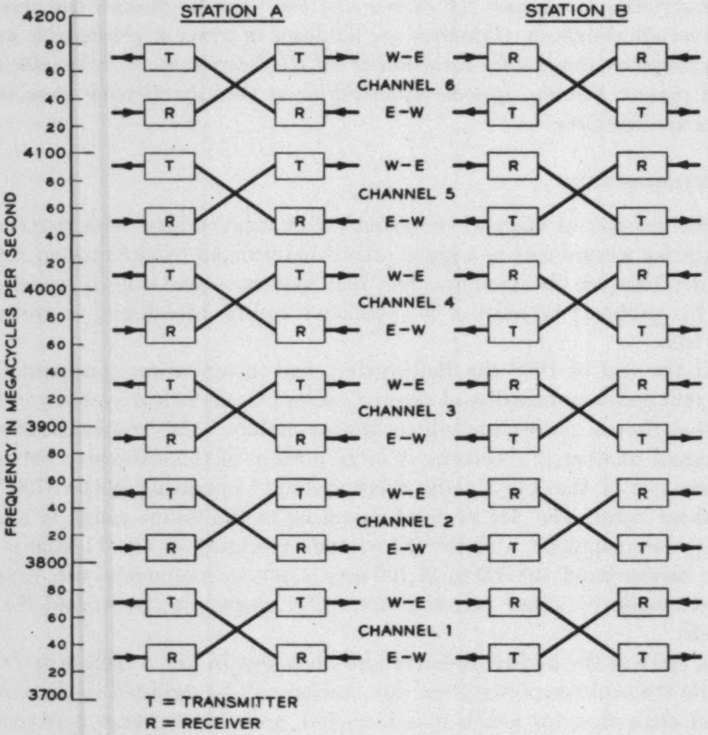
STATION A
STATION B
FREQUENCY IN MEGACYCLES PER SECOND
4200
80
60
40
20
4100
4000
3900
3800
3700
T
R
W-E
E-W
CHANNEL 6
CHANNEL 5
CHANNEL 4
CHANNEL 3
CHANNEL 2
CHANNEL 1
T = TRANSMITTER
R = RECEIVER

Fig. 1 — Frequency arrangement at a repeater.

40 mc between midband frequencies of adjacent channels in opposite directions. Thus, the potential interference from a transmitter at a tower is always separated by 40 mc or more from the closest receiving carrier frequency used at the same tower, and RF and IF filtering is sufficient to reduce such interference to a negligible magnitude.

When a two-way channel such as that shown in Fig. 1 is extended to include three consecutive repeaters, as in Fig. 2, it will be noted that two additional types of interference are involved. Similar interferences will exist, of course, on the other five channels.

For convenience in discussing these interference paths, let each antenna involved be given a number followed by a pair of letters designating the direction of transmission and whether the antenna is transmitting or receiving. Thus the sequence from west to east is 1E(T) to 2W(R) to 2E(T) to 3W(R), etc. At any particular receiving antenna such as 2W(R) there will be normally two important interferences. One is from 1W(T) into 2W(R), and the second is from 3W(T) into 2W(R). The first, for convenience, is designated "same section," since it transverses the same repeater section as the desired carrier from 1E(T); the second, 3W(T) into 2W(R), may be designated "adjacent section" for obvious reasons. Each antenna is normally subject to these same two interferences.

Of particular significance is the fact that the interference has the same nominal frequency as the carrier with which it interferes, and, hence, it cannot be discriminated against by filters. Therefore, the only protection lies in making the backward response of the antennas adequately low.

The interference situation becomes increasingly complex at a point where a number of converging microwave systems operate on the same or nearly the same frequency. An extreme example of such a point is that which exists at New York, where the Bell System has two centers from which microwave systems radiate, as shown in Fig. 3. Here we have a total of five converging routes with angles ranging from 50° to 110°, with a sixth paralleling route.

## II. RELATION BETWEEN RF INTERFERENCE AND BASEBAND INTERFERENCE

The relation between radio interference and baseband interference depends, among other things, on the kind of baseband signal being transmitted. In general, in Bell System use, this is either a video signal or a number of telephone channels arranged in frequency division. The latter case is emphasized herein, since it has been found to lead to a philosophy that provides adequate freedom from interference when either telephone or television signals are transmitted. For analytical and experimental

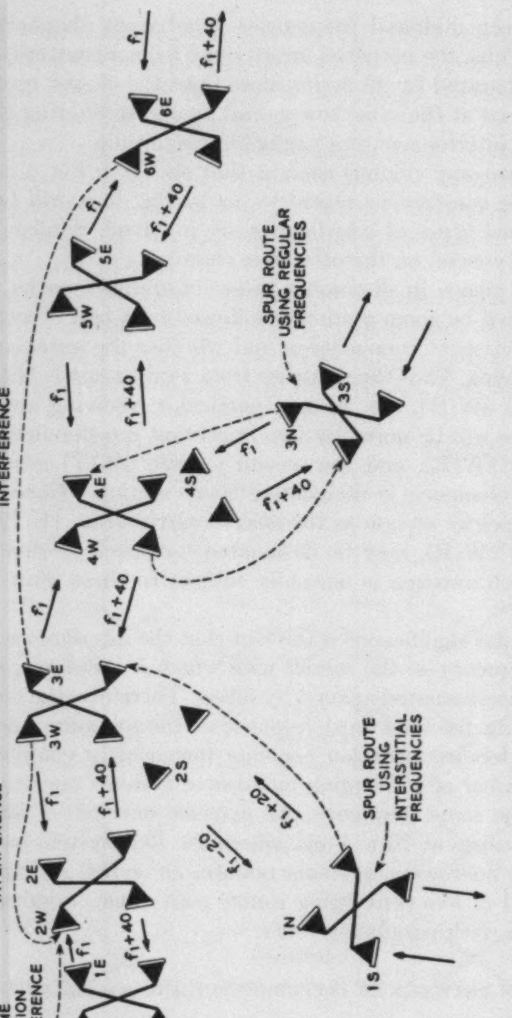
SAME SECTION INTERFERENCE
ADJACENT SECTION INTERFERENCE
OVER-REACH INTERFERENCE
SPUR ROUTE USING INTERSTITIAL FREQUENCIES
SPUR ROUTE USING REGULAR FREQUENCIES
$f_1$
$f_1 + 40$
$f_1 - 20$
$f_1 + 20$
1W
1E
2W
2E
2S
3W
3E
4W
4E
4S
5W
5E
6W
6E
1N
1S
3N
3S

Fig. 2 — Radio frequency interferences on the TD-2 microwave system with typical paths as shown.

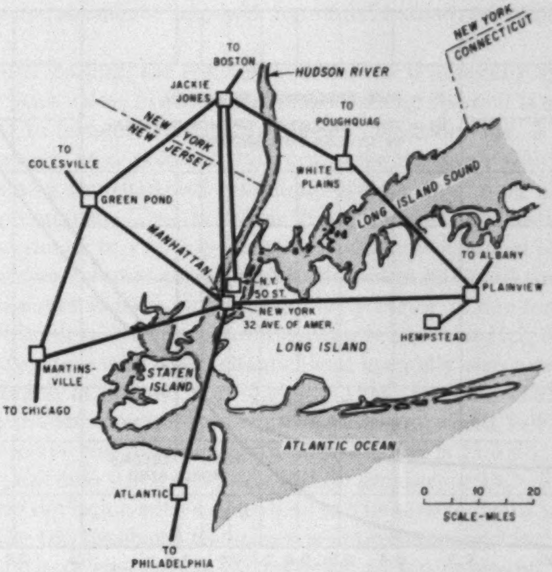
TO BOSTON
HUDSON RIVER
JACKIE JONES
NEW YORK
NEW JERSEY
TO POUGHQUAG
WHITE PLAINS
NEW YORK
CONNECTICUT
TO COLESVILLE
GREEN POND
MANHATTAN
LONG ISLAND SOUND
TO ALBANY
PLAINVIEW
N.Y. 50 ST.
NEW YORK 32 AVE. OF AMER.
LONG ISLAND
HEMPSTEAD
MARTINS-VILLE
STATEN ISLAND
TO CHICAGO
ATLANTIC OCEAN
0 10 20
SCALE-MILES
ATLANTIC
TO PHILADELPHIA

Fig. 3 — Microwave systems in New York area.

purposes a multichannel telephone signal can be adequately simulated by a band of random noise of the same average power and the same bandwidth as the multichannel signal.

The problem of intermodulation due to echoes in FM with noise loading has been treated in a previous paper,[4] which showed that the amount of intermodulation depends, among other factors, upon the echo delay.* It is of significance to the present problem that, as the echo delay becomes infinite, the correlation between the signal modulation on the echo carrier and that on the main carrier becomes zero. Hence, it may be reasoned that, in the limit, the interference problem discussed herein becomes analogous to the echo problem. An important difference is that here the carriers may be separated in frequency, whereas the echo must always have the same carrier frequency as the main wave.

This theoretical work, which is not reproduced here, leads to Fig. 4, which shows the relationship between the baseband signal-to-interference

---

* During the preparation of this paper, a theoretical paper[5] on interference between noise-modulated FM carriers appeared, which gives a general expression for the interferences as a function of frequency deviation, top baseband frequency and frequency separation between carriers. The resulting equation is identical with the one from which Fig. 4 was computed.

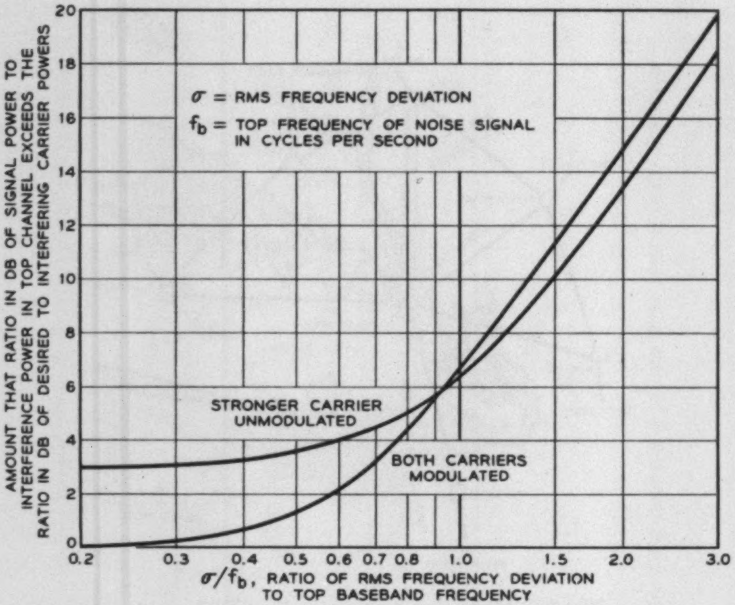

Fig. 4 — Co-channel interference from a multichannel FM system modulated with a flat band of noise.

ratio and the ratio in decibels of the desired-to-interfering carriers when the latter ratio is greater than about 10 db. This depends on the FM deviation ratio, which, in this case, is specified in terms of the ratio of the rms frequency deviation to the top baseband frequency. Furthermore, the relationship depends somewhat on whether both carriers are modulated with noise or only the weaker is so modulated. In the latter case, the baseband signal to which the interference is referred is assumed to be equal in magnitude to that applied to the interfering carrier.

In addition to the interference from the modulation on the carriers evaluated in Fig. 4, there is generated a beat note having a baseband frequency equal to the difference in frequency between the desired and interfering carriers. Since the beating oscillator frequencies in the TD-2 system are crystal-controlled, the carrier differences are held to within a few tens of kilocycles. Hence, the interferences between the carriers themselves fall below the lowest frequency used by the TD-2 system's channelizing equipment. Experience has shown that tone interference into systems carrying television service will be adequately low if the interference requirements imposed by multichannel telephone service are met.

With speech loading, the frequency deviation is generally specified in terms of the peak value, in contrast to noise loading, where it is convenient to refer to it in terms of the rms frequency deviation. Fig. 5 shows the sine wave power, at a point of 0-db transmission level, which has the same peaks as a multichannel telephone system, assuming an activity factor of one-quarter. This is taken from a paper by Holbrook and Dixon.[6] Also shown in Fig. 5 is the power of a multichannel load, based on talkers whose volumes correspond to the mean power of the assumed talker volume distribution curve, an activity factor of one-fourth being included. This curve is derived from constants given in Ref. 6.

The TD-2 system with a 480-channel load normally uses a peak deviation with speech of ±4 mc. Fig. 5 shows that, for this channel load, the ratio of the rms frequency deviation with equivalent full sine wave modulation to the rms deviation with noise loading is 11.5 db. This ratio referred to peak sine wave deviation is 3 db greater, or 14.5 db. Accordingly, the rms deviation with a noise load can be seen to be 0.7 mc. Therefore, since the top baseband frequency is approximately 2 mc in the example, the ratio of rms frequency deviation to top baseband frequency would be approximately 0.35. Hence, Fig. 4 shows that, in this case, the baseband signal-to-interference ratio in a telephone channel at the top

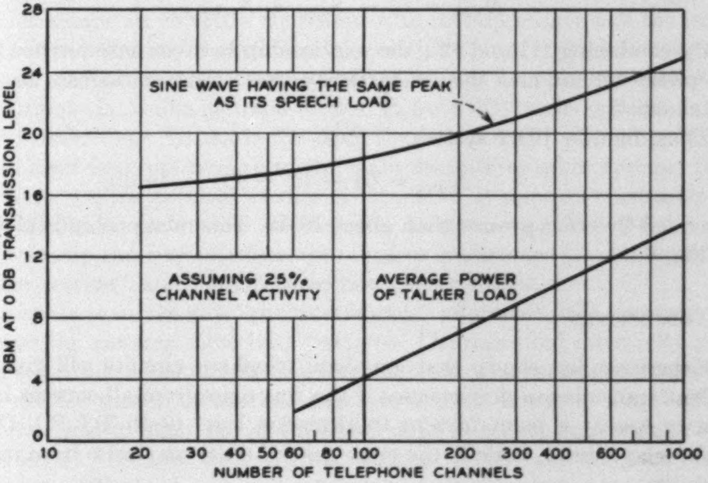

Fig. 5 — Multichannel telephone load, assuming 25 per cent channel activity.

baseband frequency is $\frac{1}{2}$ db greater than the ratio in decibels of the "desired" carrier to an interfering carrier when both carriers are modulated with noise. Thus, by rounding this value to 1 db, we may write, in db:

$$S_{\text{dbm}} - N_{\text{dbm}} = C - I + 1. \tag{1}$$

It has been found useful in the Bell System to express noise power in a telephone circuit in dba.* This implies a specific frequency weighting characteristic, and 1 milliwatt (0 dbm) of flat noise in a 3-kc band reads 82 dba. For a measurement of noise power to have meaning, it must be referred to the signal power at the same point in the circuit, or else the transmission level at which the measurement is made must be specified. Noise in dba is often specified at the 0-db transmission level, the point of reference being the toll transmitting switchboard.

Holbrook and Dixon give the power of an rms talker at a point of 0-db transmission level (0-db T.L.P.) as $-9.9$ dbm. With a channel activity factor of one-quarter, the equivalent noise load per 4-kc channel is $-15.9$ dbm. The noise in a 3-kc band is 1.2 db less, or $-17.1$ dbm. This amount of noise power would measure very nearly, 65 dba. From this it follows that, given a measurement of signal power and of noise power in the same narrow band, the equivalent noise in dba at a point of zero transmission level can be obtained from the following relationship:

$$N_{\text{dba}} = 65 - S_{\text{dbm}} + N_{\text{dbm}}. \tag{2}$$

By combining (1) and (2), the relationship between interference in a telephone circuit and the RF carrier-to-interference ratio can also be obtained.

Thus, for the TD-2 system,

$$N_{\text{dba}} = 64 - (C - I), \qquad \text{in db,} \tag{3}$$

provided $C - I$ is greater than about 10 db. This relationship is plotted in Fig. 6.

## III. OBJECTIVES

Experience has shown that long-haul telephone circuits will give excellent transmission performance if the rms noise from all sources is 38 dba or less at a point of zero transmission level (0-db T.L.P.). On a radio relay circuit, noise in the more general sense may arise from many

---

* A discussion of this unit and its relation to other units of noise power is given in some detail by Franke.[7]

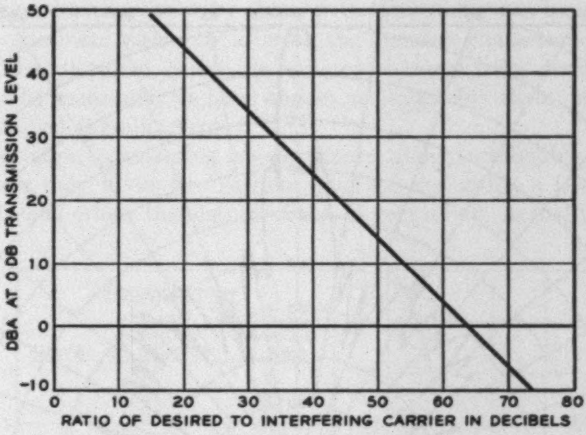

Fig. 6 — Telephone channel interference due to an interfering co-frequency radio channel in TD-2 and TH microwave systems, all-telephone case.

sources. A portion will be in the form of fluctuation noise arising in the converters; some will be intermodulation due to amplitude and phase distortion in the IF and RF circuits, as well as from echoes in the waveguide runs and the interconnecting IF cables. It is necessary to divide the total permissible noise among all the known contributors, alloting to each source an objective consistent with its importance and the feasibility of meeting it. Of particular concern here is unintelligible crosstalk due to RF interference; a reasonable allocation for this last source of impairment is 24 dba at the 0-db T.L.P. for a 4000-mile system.

Provision must be made for both the interference within the main route itself and interference to the main route from other systems that converge. Of the total of 24 dba given above, it appears reasonable to allot 22 dba to main route interference. As mentioned above, there will be two main-route RF interferences of major importance at each repeater: "same section" and "adjacent section" interference.

In order to arrive at a "per interference" allotment, it is necessary to discuss the antenna directivity patterns. The measured directivity pattern of an individual sample of a given antenna at a particular frequency will exhibit numerous maxima and minima, as shown by Fig. 7. It is not practical to make use of the sharp nulls to reduce converging route interference, because their angular position varies with frequency and is apt to differ from one particular antenna to another. Also, the nulls are usually sharper than the angular stability of the tower used to sup-

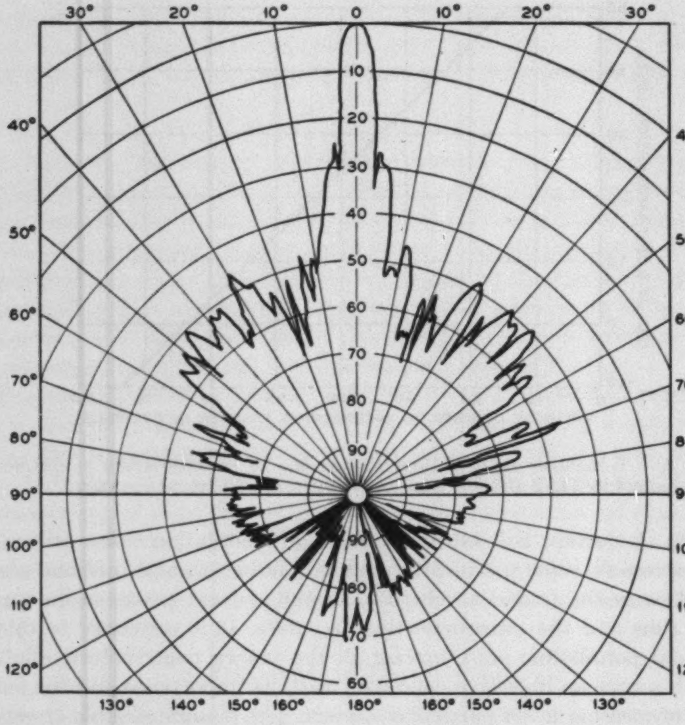

Fig. 7 — Measured directivity pattern of an individual antenna (horn-reflector, horizontal directivity, vertical polarization, 3740 mc, 3-db beamwidth 2.5°).

port the antenna. Therefore it is desirable to use smoothed directivity patterns, as shown in Figs. 8, 9 and 10, in designing particular systems. The smoothed patterns were obtained by drawing a more or less smooth curve through the peak (or poorest) values of the measured patterns. Therefore, these effectively are "worst" values rather than "rms" values.

In a multirepeatered system where there is a sizable number of exposures involving random angles, a substantial percentage of these exposures will contribute less interference than would be indicated by the work presented here. Therefore, the total rms interference is somewhat less than would be expected by simple power addition of the individual interferences. An examination of various measured directivity patterns indicates that it is reasonable to assume that straight power addition is perhaps too conservative by about 6 db. This factor can be introduced into the allotment either by altering the antenna characteristics or by altering the allotment. Since the antenna characteristics describe what was actually measured, we have chosen not to modify them; instead, we have modified the allotment.

Furthermore, considering one exposure alone, intelligible crosstalk can appear only if the two carriers involved are within a few hundred cycles of each other; the chance of this is very small. Actually, the two

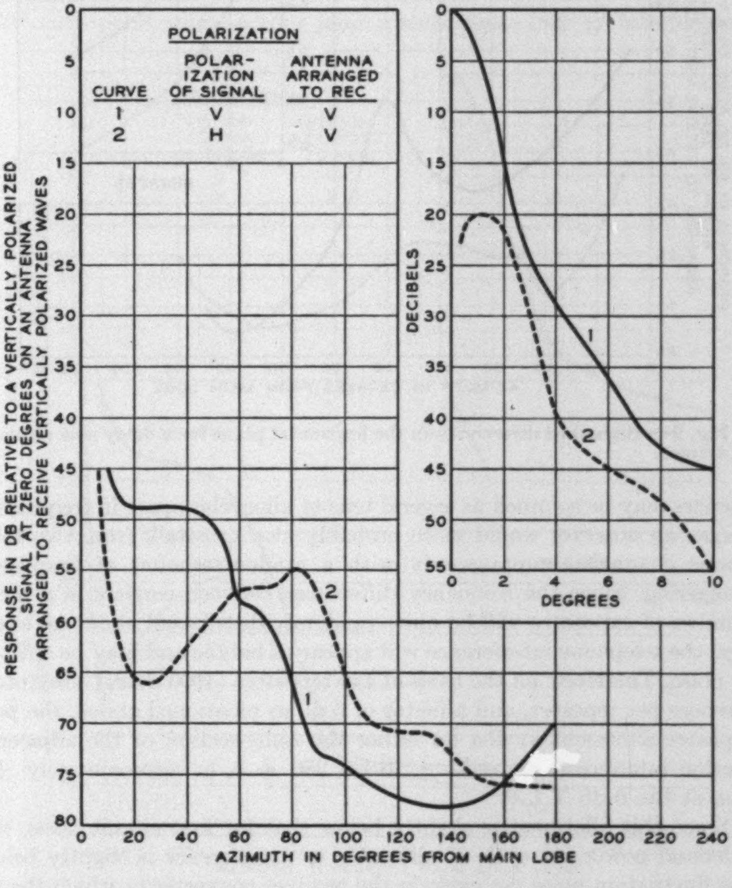

Fig. 8 — Smoothed directivity in the horizontal plane for a horn-reflector antenna at 4 kmc.

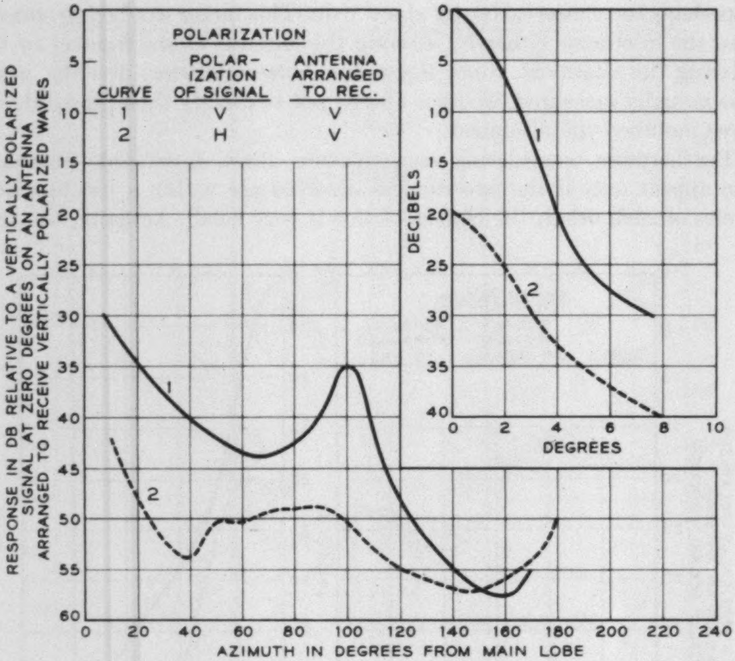

Fig. 9 — Smoothed directivity in the horizontal plane for a delay lens antenna at 4 kmc.

carriers may be as much as several tens of kilocycles apart in frequency; hence, an observer would more probably hear crosstalk from two telephone channels simultaneously with a random amount of frequency staggering. Since the frequency differences between carriers in a large number of exposures will be quite randomly distributed about an average, the telephone interference will appear as babble and may be treated as noise. Therefore, on the basis of 140 repeaters (4000 miles), two interferences per repeater, and a factor of 6 db as mentioned above, the per-repeater allotment in dba for either the same-section or the adjacent-section interference is: $+22 - 10 \log 280 + 6$, or approximately $+4$ dba at the 0-db T.L.P.

Since this allotment is slightly below that for first circuit noise, the sideband power per cycle of allowable RF interference is slightly below the fluctuation noise per cycle in the receiver converter in which the RF interference falls.

It is then evident that, during a fade of the "desired" carrier entering any particular receiver converter, the converter noise and the RF interference contributions, as observed in any telephone channel at the terminal of the system, will rise together and always stay in balance. Obviously, if the interfering carrier fades it becomes increasingly weaker than the fluctuation noise, and the effect then becomes entirely negligible.

Pertinent features of a long main route such as may extend across the continent are the spurs by means of which television and telephone channels are brought to cities situated off the main route.

At each point where a spur joins a main route four RF interferences

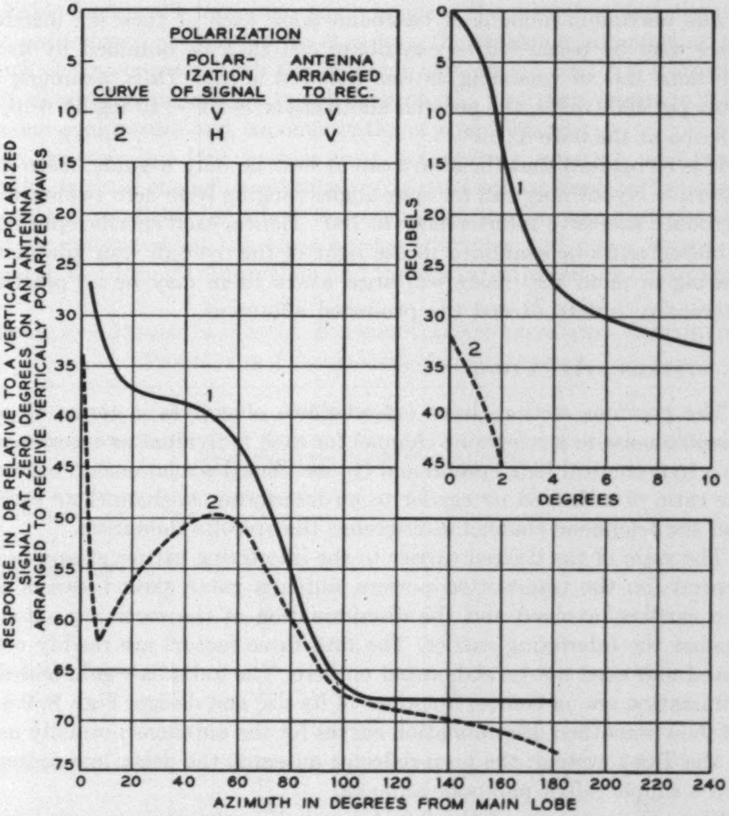

Fig. 10 — Smoothed directivity in the horizontal plane for an 8-foot parabolic antenna at 4 kmc.

will occur, as shown in Fig. 2. Two of these will appear finally as telephone channel babble at the west end of the backbone route, and similarly two will appear at the east end. Thus, assuming 64 spurs per 4000 miles, there will be a maximum of 128 such interferences at either end.

However, the angles of the spurs may be expected to be distributed at random, with presumably no angle less than a value consistent with the objectives developed below. Hence, many interferences may be expected to contribute substantially less baseband noise than the alloted value. In view of this it is probably not unreasonable to assume that only one-half of these interferences will contribute significantly to the total.

The maximum amount of baseband noise each of these RF interferences may be permitted to contribute can then be obtained by using the same line of reasoning as was followed above. Thus, assuming 64 spurs per 4000 miles, the per spur allotment is: $+20 - 10 \log 64 + 6$, or $+8$ dba at the 0-db T.L.P.

It is recognized that the above can at best be only a guide, since any tentative layout may call for spur angles ranging from zero (which may introduce excessive interference) to 180°. Hence, each specific spur contribution must be examined in the light of the over-all spur allotment, bearing in mind that cases will arise where there may be no practical alternative but to exceed the pro-rated allotment.

## IV. SYSTEMS APPLICATION

The previous sections have (a) set down objectives in terms of permissible noise in a telephone channel for each individual RF interference on a hypothetical long system and (b) developed a relationship between the ratio of a desired RF carrier to an interfering co-channel RF carrier and the telephone channel interference that results therefrom.

The ratio of the desired carrier to the interfering carrier at any point depends on the transmitted powers, antenna gains, path losses of the two carriers involved and the discrimination of the receiving antenna against the interfering carrier. The first three factors are readily computed and need not be elaborated on here. The antenna's gain and discrimination are, of course, functions of its size and design. Figs. 8, 9 and 10 show smoothed discrimination curves for the antennas normally used in the TD-2 system: the horn-reflector antenna, the delay lens antenna and a simple 8-foot parabolic antenna.

Since the response of the 8-foot parabolic antenna to a signal from the rear is only 45 db below the main lobe response, both the "adjacent section" and "same section" interference would be on the average +19 dba, or a total of +22 dba per repeater. This would be regarded as intolerably high for a long-haul circuit. For this reason, the parabolic antennas have been used mostly on short-haul spur or secondary routes. Both the delay lens and horn-reflector provide in excess of 65 db discrimination in the rearward direction, which is entirely adequate even for systems of 4000 miles in length.

The minimum angle between a spur route and a main route on the basis of a given permissible interference is determined by the discrimination of the antennas involved. This, for a given pair of antennas, depends on whether the interference is polarized similarly or at right angles to the plane of polarization of the desired carrier.

Table I gives the minimum angle between the disturbed and disturbing paths based on the interference into the receiving antennas at the point of convergence, with the inherent assumption that the disturbed and disturbing signals are equal in strength; i.e., they originate from similar equipments and traverse paths of equal length.

In general, the Bell System does not operate separate microwave routes in close enough proximity for mutual interference to become a consideration. However, where it is suspected, the baseband interference can be computed on the basis of the principles laid down above. In general, disturbing transmitters and the disturbed receivers on separate

TABLE I — MINIMUM ANGLE BETWEEN DISTURBED AND DISTURBING PATHS FOR CO-CHANNEL INTERFERENCE

| Disturbed Antenna Type | Polarization | | Minimum Angle (Degrees) |
|---|---|---|---|
| | Disturbing (Transmitting) Antenna | Disturbed (Receiving) Antenna | |
| Delay lens | V | V | 80 |
| | H | V | 72 |
| Horn-reflector | V | V | 58 |
| | H | H | 26 |
| | H | V | 10 |
| | V | H | 11 |
| Parabolic, 8-foot | V | V | 144 |
| | H | H | * |
| | H | V | 130 |
| | V | H | 26 |
| Parabolic, 10-foot | V | V | 122 |
| | H | H | * |
| | H | V | 110 |
| | V | H | * |

* The objective cannot be met at any angle with these conditions.

routes do not point at one another, and the discrimination advantage of both antennas is obtained. The combinations and possibilities are so great for such cases that generalization becomes useless and each case should be treated individually.

In the standard frequency arrangement for the TD-2 system, transmitting and receiving carriers are separated by 40 mc. Ahead of the receiver converter, RF filters provide 20 db of discrimination to the transmitter. This and the IF selectivity are more than adequate to eliminate interference between transmitters and receivers at the same site.

It is possible under certain conditions to operate channels in the 20-mc slots between an adjacent transmitter and receiver in the standard frequency arrangement. These interstitial or so-called "slot" channels are used at converging points or spurs where co-channel interference considerations would limit the minimum angle of convergence to an undesirably large angle.*

When interstitial frequencies are used, each channel may have interferences separated only 20 mc from the normal carrier frequency, i.e., at 50 mc and 90 mc in the IF band. This separation is not sufficient to provide any appreciable suppression of the interfering carriers by RF filtering. Furthermore, the IF discrimination to such interferences is only of the order of 5 to 10 db.

If the FM receiver could be assumed to consist of a perfect limiter and a perfect demodulator, FM theory would indicate that the resulting interference at baseband frequencies would be infinitesimal in magnitude in the frequency range occupied by the normal signal. Furthermore, any such interference would be incoherent with the modulating signal on the disturbing channel. Actually, experiment has shown that the signal on the weaker carrier is transferred to the stronger carrier. The crosstalk increases 2 db as the ratio between the desired carrier and the interfering carrier decreases 1 db, provided that the interfering carrier is weaker than the desired carrier; beyond this point, the baseband interference rises still more rapidly. No completely satisfactory explanation has been advanced, other than the generalized one that physical amplifiers, limiters and discriminators are frequently not ideal devices.

---

* Developments subsequent to the preparation of this paper indicate that additional discrimination applied at each repeater at the intermediate frequency of 70 mc reduces the crosstalk from the adjacent microwave channels 20 mc offside so as to allow the minimum angle between converging routes to become zero degrees. This is equivalent to saying that six interstitial two-way microwave channels can also be operated in the 3700- to 4200-mc band, making a total of 12 broadband channels. The technical problems relating to the possible use of these six additional two-way channels are now being studied at Bell Telephone Laboratories.

From a systems engineering viewpoint, when interstitial frequencies are used the interfering carrier must be sufficiently weaker than the desired carrier during free-space transmission that, when the desired carrier fades, the crosstalk remains below the fluctuation noise as long as the circuit is useful.

Fig. 11 shows how the fluctuation noise theoretically increases with depth of fade on a single typical repeater section. Also in Fig. 11 is a curve showing how the crosstalk from an interstitial channel increases

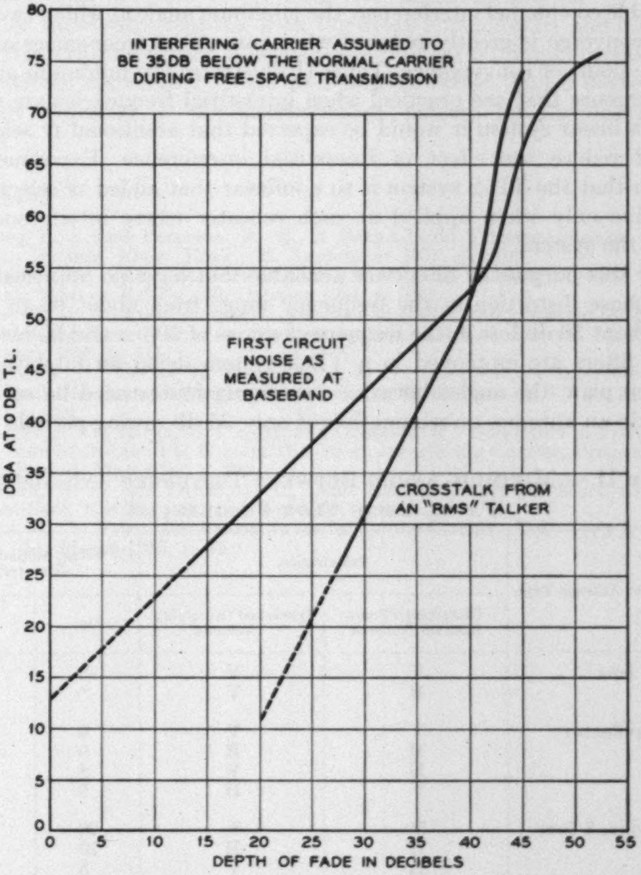

Fig. 11 — First circuit noise and adjacent channel crosstalk, with interfering carrier assumed to be 35 db below normal carrier during free-space transmission.

with a fade, assuming that the interfering carrier is 35 db weaker than the interfered-with carrier during periods of free-space transmission. It is assumed, in this case, that the interfering channel is modulated with average power talkers.

This figure indicates that, with a carrier ratio of 35 db, the crosstalk remains below the noise for fades up to about 35 db. Arrangements are provided automatically to switch the faded channel to a protection channel when the depth of fade exceeds 35 db.

Since this carrier ratio is about 20 db less than the requirement for tolerable co-channel interference, the minimum angle at which two routes may converge is greatly reduced when interstitial frequencies are used at the point of convergence. Table II illustrates the minimum angles of convergence that are practical when interstitial frequencies are used.

In a linear system it would be expected that additional IF selectivity would reduce the effect of interstitial interference. Experiment has shown that the TD-2 system is so nonlinear that added IF selectivity is effective only when applied at each repeater where interference may enter the system.

For this purpose IF filters are available that have no appreciable loss and phase distortion in the frequency range from about 60 to 80 mc, and about 30 db loss in the frequency regions of 50 mc and 90 mc. When these filters are employed in a TD-2 system using an interstitial frequency plan, the angles between converging routes need be such as to provide an antenna discrimination of only 25 db during periods of free-

TABLE II — MINIMUM ANGLE BETWEEN DISTURBED AND DISTURBING PATHS USING SLOT FREQUENCIES

| Disturbed Antenna Type | Polarization | | IF Filter, Minimum Angle (Degrees) | |
|---|---|---|---|---|
| | Disturbing (Transmitting) Antenna | Disturbed (Receiving) Antenna | Out | In |
| Delay lens | V | V | 12 | 3.5 |
| | H | V | * | * |
| Horn-reflector | V | V | 6 | 3.1 |
| | H | H | 9 | 5.5 |
| | H | V | 4 | 2.5 |
| | V | H | 6 | 2.4 |
| Parabolic, 8-foot | V | V | 20 | 5 |
| | H | H | 20 | 7.2 |
| | H | V | 5 | 1.8 |
| | V | H | 8 | 4.9 |

* Less than 2°; a minimum of 2° is suggested.

space transmission. Table II gives the minimum angles necessary for adequate protection from interference when IF filters are used.

When the suggested minimum angles are very small, as when filters are used, allowance must be made for the mechanical stability of the towers and for the precision to which the antennas can be oriented. The former depends on the design of the tower structure and on the loads to which it is subjected. The latter depends on the methods and skill used in orienting the antennas. No safety factors for these considerations have been included in the suggested minimum angles.

Also for these same reasons, with the delay lens antenna and a cross-polarized interfering carrier, it is suggested that the minimum permissible angle between converging routes be 2 degrees rather than 0 degrees, as would be indicated by the discrimination pattern for the delay lens antenna.

## REFERENCES

1. Grieser, T. J. and Peterson, A. C., A Broad-Band Transcontinental Radio Relay System, Elect. Engg., **70,** September 1951, p. 810.
2. Roetken, A. A., Smith, K. D. and Friis, R. W., The TD-2 Microwave Radio Relay, System, B.S.T.J., **30,** October 1951, p. 1041.
3. McDavitt, M. B., 6,000-Megacycle-Per-Second Radio Relay System for Broad-Band Long-Haul Service in the Bell System, A.I.E.E. Trans., **76,** Part 1, January 1958, p. 715.
4. Bennett, W. R., Curtis, H. E. and Rice, S. O., Interchannel Interference in FM and PM Systems Under Noise Loading Conditions, B.S.T.J., **34,** May 1955, p. 601.
5. Medhurst, R. G., Hicks, Mrs. E. M. and Grossett, W., Distortion in Frequency Division Multiplex FM System Due to an Interfering Carrier, Proc. I.E.E., **105,** Part B, May 1958, p. 282.
6. Holbrook, B. D. and Dixon, J. T., Load Rating Theory for Multi-Channel Amplifiers, B.S.T.J., **18,** October 1939, p. 624.
7. Franke, H. C., Noise Measurements on Telephone Circuits, Tele-Tech & Elect. Ind., **14,** March 1955, p. 85.

# Diffused Junction Depletion Layer Calculations

By H. LAWRENCE and R. M. WARNER, Jr.

(Manuscript received October 7, 1959)

*Depletion layer properties have been calculated for diffused junctions in silicon and germanium as a function of reverse voltage and of diffusion parameters for the gaussian and the complementary error function distributions. These results bridge the gap between the linearly graded behavior generally exhibited by such junctions at low voltage and the step behavior exhibited at high voltage. For total depletion layer thickness and capacitance, the transition from graded to step junction behavior extends over about one decade of voltage. For depletion layer thickness on a single side of the junction, it extends over several decades. Depletion layer thickness and peak electric field are presented graphically as a function of voltage for a variety of junction depths and impurity concentration functions. The ranges for which the step and graded junction approximations are valid are apparent from these charts. The results were obtained by an analytical integration of Poisson's equation, and a subsequent use of the IBM 704 for a numerical evaluation of the transcendental equations obtained.*

## I. INTRODUCTION

The dependence of depletion layer properties on voltage is important in the design of many semiconductor devices. Some of these properties, such as total depletion layer thickness, are accurately predicted by the commonly used step and graded junction approximations over fairly wide voltage ranges. However, the ranges in which the approximations are applicable have not been established previously. The present work establishes the regions in which these approximations are valid and supplies data for the entire voltage range of interest.

Both the complementary error function distribution and the gaussian distribution have been treated. Total depletion layer thickness, peak electric field, capacitance per unit area and the fraction of the depletion layer on each side of the junction have been calculated for wide ranges of voltage, junction depth, background impurity concentration and surface concentration. Results are given for silicon and germanium and can readily be extended to other materials.

## II. METHOD — COMPLEMENTARY ERROR FUNCTION DISTRIBUTION

For the case of diffusion into a semiconductor with the assumption of a concentration-independent diffusion constant and a constant volume-concentration at the surface, the solution of the one-dimensional diffusion equation is the complementary error function.[1] The solution may be written

$$C(x) = C_0 \operatorname{erfc} \frac{x}{\sqrt{4Dt}}, \tag{1}$$

where $C$ is the impurity density at a distance $x$ from the surface, $C_0$ the surface concentration, $D$ the temperature-dependent diffusion constant and $t$ the time of diffusion.*

For the purpose of these calculations we have taken the case of a donor diffusion into a p-type semiconductor. To treat the depletion layer formed at the junction, when reverse bias is applied, we have made the following customary assumptions:

i. ionization of donors and acceptors is complete;

ii. the depletion layer is a region completely free of carriers separated by a sharp boundary from an electrically neutral region.

The net density of positive charge, $\rho$, lying within the depletion layer can be written

$$\rho = q(C - C_B), \tag{2}$$

where $q$ is the electronic charge and $C_B$ is the background impurity concentration. Hence Poisson's equation can be written in MKS units as follows:

$$\nabla^2\psi = -\frac{q}{\kappa\epsilon_0}\left(C_0 \operatorname{erfc} \frac{x}{\sqrt{4Dt}} - C_B\right), \tag{3}$$

where $\psi$ is electrostatic potential, $\kappa$ is the dielectric constant of the material and $\epsilon_0$ is the permittivity of free space.

As Fig. 1 indicates, the electric field vanishes at the depletion layer boundaries. Thus, by integrating Poisson's equation, we can write expressions for the field distribution on the two sides of the junction: On the left,

---

* $\operatorname{erfc} u = 1 - \operatorname{erf} u; \quad \operatorname{erf} u = \frac{2}{\sqrt{\pi}} \int_0^u \exp(-\alpha^2)\, d\alpha.$

$$\begin{aligned} E_1(x) - E_1(x_j - a_1) &= E_1(x) \\ &= \frac{qC_0}{\kappa\epsilon_0}\int_{x_j-a_1}^{x}\left(1 - \frac{C_B}{C_0} - \operatorname{erf}\frac{x}{\sqrt{4Dt}}\right)dx \end{aligned} \tag{4}$$

and, on the right,

$$\begin{aligned} E_2(x_j + a_2) - E_2(x) &= -E_2(x) \\ &= \frac{qC_0}{\kappa\epsilon_0}\int_{x}^{x_j+a_2}\left(1 - \frac{C_B}{C_0} - \operatorname{erf}\frac{x}{\sqrt{4Dt}}\right)dx, \end{aligned} \tag{5}$$

where $x_j$, the junction depth, is defined by

$$C_0 \operatorname{erfc}\frac{x_j}{\sqrt{4Dt}} - C_B = 0. \tag{6}$$

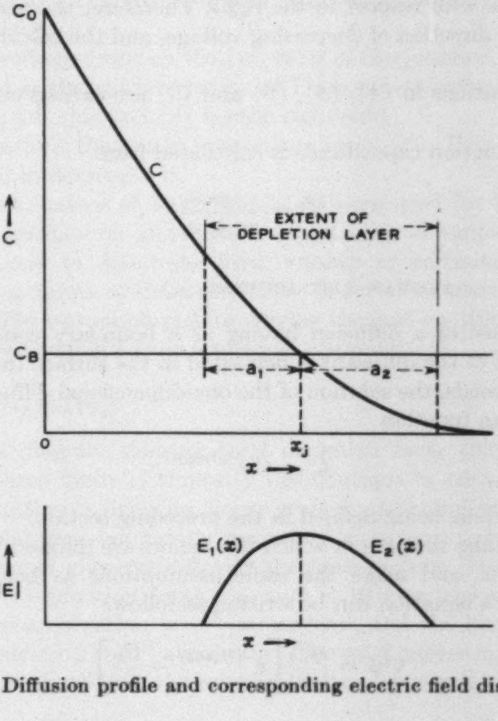

Fig. 1 — Diffusion profile and corresponding electric field distribution.

Requiring that the field be continuous at the junction enables us to arrive at the desired relation between the two components of depletion layer thickness, $a_1$ and $a_2$. That is,

$$E_1\left(x_j, \frac{C_B}{C_0}, a_1\right) = E_2\left(x_j, \frac{C_B}{C_0}, a_2\right). \tag{7}$$

At this point specific values can be chosen for $x_j$, $C_B/C_0$ and $a_1$ so that $a_2$ can be calculated by a machine using iterative methods.

Next, the voltage drops on the two sides of the junction are obtained by a second analytical integration of Poisson's equation:

$$V_1 = -\int_{x_j-a_1}^{x_j} E_1(x)\,dx, \tag{8}$$

$$V_2 = -\int_{x_j}^{x_j+a_2} E_2(x)\,dx. \tag{9}$$

For the situation chosen, reverse bias implies a positive voltage on the left-hand side with respect to the right. Therefore, the positive $x$ direction is in the direction of decreasing voltage, and the calculated voltages are negative.

The integrations in (4), (5), (8) and (9) are carried out in detail in Appendix A.

Finally, junction capacitance is calculated from

$$C = \frac{\kappa\epsilon_0}{a_1 + a_2}. \tag{10}$$

## III. METHOD — GAUSSIAN DISTRIBUTION

For the case of a diffusion having as a boundary condition that a fixed amount of the diffusant is deposited in the surface from which the diffusion proceeds, the solution of the one-dimensional diffusion equation is the gaussian function

$$C = C_0 e^{-x^2/(4Dt)}, \tag{11}$$

with the symbols being defined in the preceding section.

Again we take the case in which the donors are diffused into a p-type semiconductor, and make the same assumptions as before. For this case, Poisson's equation can be written as follows:

$$\nabla^2\psi = \frac{qC_0}{\kappa\epsilon_0}\left[e^{-x^2/(4Dt)} - \frac{C_B}{C_0}\right]. \tag{12}$$

By integrating Poisson's equation as before we can write expressions for the field distribution on the two sides of the junction:
On the left,

$$E_1(x) - E_1(x_j - a_1) = E_1(x) = \frac{qC_0}{\kappa\epsilon_0}\int_{x_j-a_1}^{x}\left[e^{-x^2/(4Dt)} - \frac{C_B}{C_0}\right]dx \quad (13)$$

and, on the right,

$$E_2(x_j + a_2) - E_2(x) = -E_2(x) = \frac{qC_0}{\kappa\epsilon_0}\int_{x}^{x_j+a_2}\left[e^{-x^2/(4Dt)} - \frac{C_B}{C_0}\right]dx, \quad (14)$$

where $x_j$, the junction depth, is defined by

$$[C_0e^{-x_j^2/(4Dt)} - C_B] = 0. \quad (15)$$

Requiring that the field be continuous at the junction enables us to arrive once again at the desired relation between the two components of depletion layer thickness, $a_1$ and $a_2$. This equation, (7), is treated by machine as before.

Next, the voltage drops on the two sides of the junction are obtained by the second analytical integration of Poisson's equation, as indicated in (8) for the left side and (9) for the right side.

The integrations that yield expressions for $E_1(x)$, $E_2(x)$, $V_1$ and $V_2$ are completed in Appendix B.

In this work, values of 16.00 and 12.00 were used for the dielectric constants of germanium and silicon respectively. A simple multiplication enables one to determine field, voltage or capacitance for any modification of values of these constants, or for other materials.

The IBM 704 was employed for solving the field continuity equation (7) and for performing all the numerical computations.

## IV. DATA PRESENTATION

Fig. 2 is a diagram showing total depletion layer thickness versus voltage for three kinds of impurity distributions in silicon: step with large concentration on one side, linear graded, and complementary error function. Using gaussian results instead of complementary error function results would give qualitatively the same picture. The relationship is shown for two junction depths: $x_j = 2 \times 10^{-3}$ cm and $x_j = 10^{-4}$ cm. The surface concentration is $10^{20}$ atoms/cm$^3$, and the background impurity concentration is $10^{16}$ atoms/cm$^3$. The solid curves marked "erfc," which were obtained from the calculated data, clearly mark the transi-

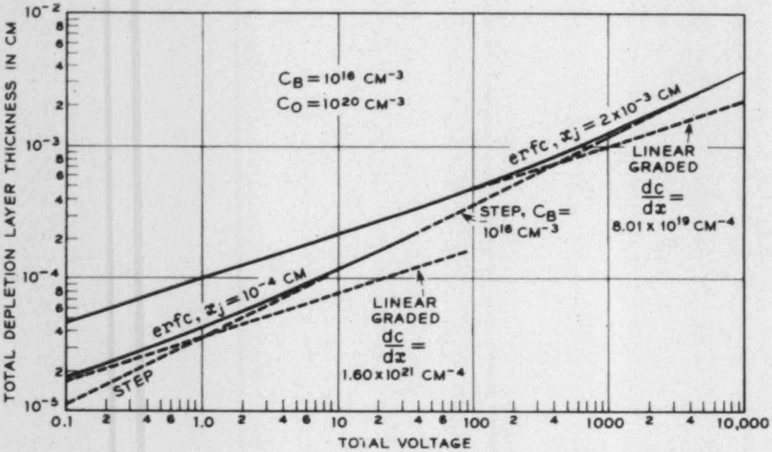

Fig. 2 — Total depletion layer thickness vs. voltage for step, linear graded and complementary error function impurity distributions in silicon.

tion region between the step and linear distributions, becoming asymptotic to these two curves at the higher and lower voltages respectively. It should be noted that, since the step junction curve is a function of background resistivity only, it remains the same for both junction depths.

The intersection of the linear graded and the step junction lines is the point where the accuracy of either approximation is poorest. Increasingly larger errors occur, of course, if either approximation is used for voltages beyond its point of intersection with the other approximation. The curves shown here constitute extreme examples. For $x_j = 10^{-4}$ cm the step junction approximation becomes excellent at about 4 volts, and hence the step approximation can be used over almost the entire voltage range of interest. In the case of $x_j = 2 \times 10^{-3}$ cm, however, breakdown occurs around 100 volts, and therefore the graded approximation is good over almost the entire voltage range of interest.

Fig. 3 presents total depletion layer thickness versus the ratio of voltage to background impurity concentration. Fig. 4 presents the corresponding data for the gaussian distribution. These charts are available for a wide range of values of the ratio $C_B/C_0$ ; a complete set will be furnished by the authors upon request.

The chart in Fig. 3 is precise for $C_B/C_0 = 10^{-5}$, but one may use the chart for a range from $3 \times 10^{-6}$ to $3 \times 10^{-5}$. Maximum errors resulting

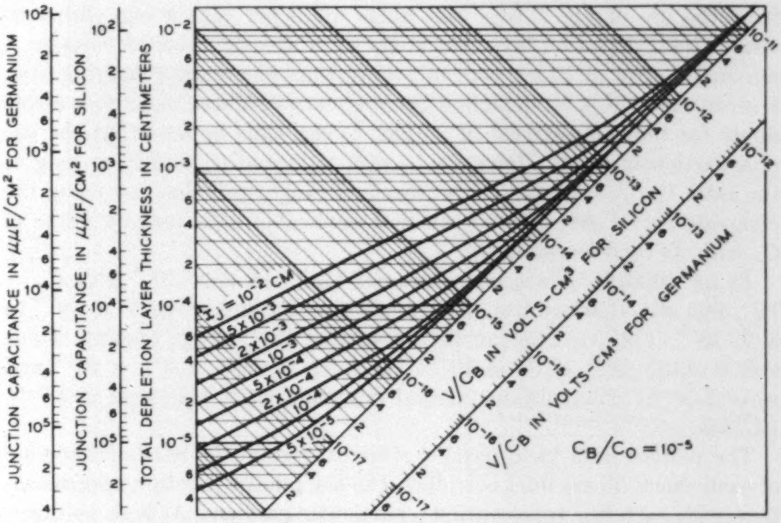

Fig. 3 — Chart for use in range $3 \times 10^{-6}$ to $3 \times 10^{-5}$, erfc distribution.

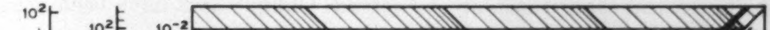

Fig. 4 — Chart for use in range $3 \times 10^{-6}$ to $3 \times 10^{-5}$, gaussian distribution.

from the use of Fig. 3 over this range are 5 per cent if one wishes to obtain total depletion layer thickness for a specific value of voltage. A maximum error of 10 per cent in voltage may result when starting from a given value of total depletion layer thickness. These maximum errors apply for the complete set of graphs. It can thus be seen that the use of a particular graph allows one to shift values of $C_B$ and $C_0$ as long as the ratio $C_B/C_0$ falls within the specified range, and that, by using the complete set of graphs, one can find applicable data for any value of $C_B$ with its corresponding $C_0$.

As an example, if one has a ratio of $C_B/C_0 = 2 \times 10^{-5}$ with $C_B = 10^{15}$, one would select the chart applicable for the range $3 \times 10^{-6}$ to $3 \times 10^{-5}$ (Fig. 3). One can then note that the voltage reading of one volt is obtained at $V/C_B = 10^{-15}$. Then, for 10 volts, $V/C_B = 10^{-14}$ and $x_j = 2 \times 10^{-3}$, one obtains a total depletion layer thickness of $4.7 \times 10^{-4}$ cm.

The plotted data yield parallel straight lines in the left-hand portion of each chart. These lines correspond to the graded junction approximation, with each line representing a particular gradient. At high voltages, the lines then converge onto the single straight line that corresponds to the step junction approximation. Thus, these charts define clearly the applicability ranges of the two approximations insofar as total layer thickness and capacitance are concerned. Oblique axes have been used to spread the curves conveniently.

Since $V/C_B$ is used as a parameter in the charts, avalanche breakdown data cannot be readily superimposed. The problem can be appreciated by noting that breakdown voltage depends on background doping in a step junction, but all step junction data have been collapsed onto a single line in these charts.

Fig. 5 shows the variation of $a_1/a_{\text{total}}$ with the voltage function for the complementary error function distribution. Fig. 6 gives the corresponding curves for the gaussian function. The upper boundary of each chart corresponds to the graded junction, and the lower boundary corresponds to the step junction. Note that here the transition range covers many decades.

A plot of the peak electric field divided by $C_B$ versus $V/C_B$ is given in Fig. 7. This chart was made for $C_B/C_0 = 10^{-6}$, for the case of the complementary error function. The usefulness of the chart can be extended to include a range of $C_B/C_0$ from $10^{-4}$ to $10^{-8}$, where the maximum error for a $C_B/C_0$ other than $10^{-6}$ is 10 per cent and, in the majority of cases, the error would be less than 5 per cent. For the gaussian distribution, the usable range of this same graph is from $C_B/C_0 = 10^{-4}$ to $C_B/$

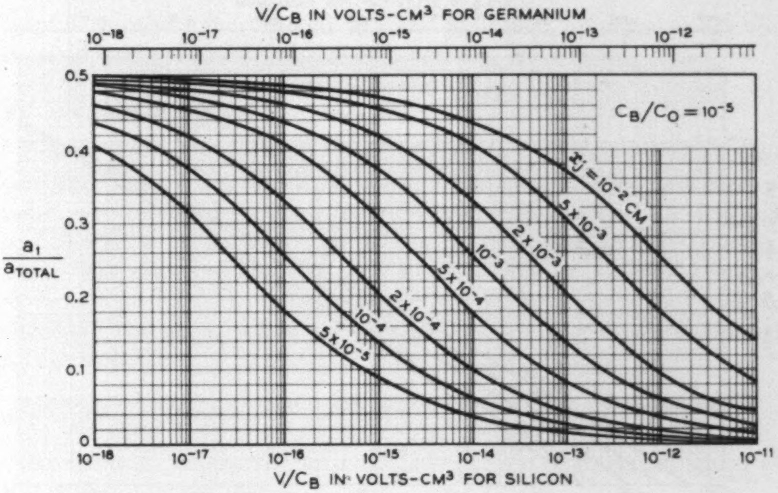

Fig. 5 — Chart for use in range $3 \times 10^{-6}$ to $3 \times 10^{-5}$, erfc distribution.

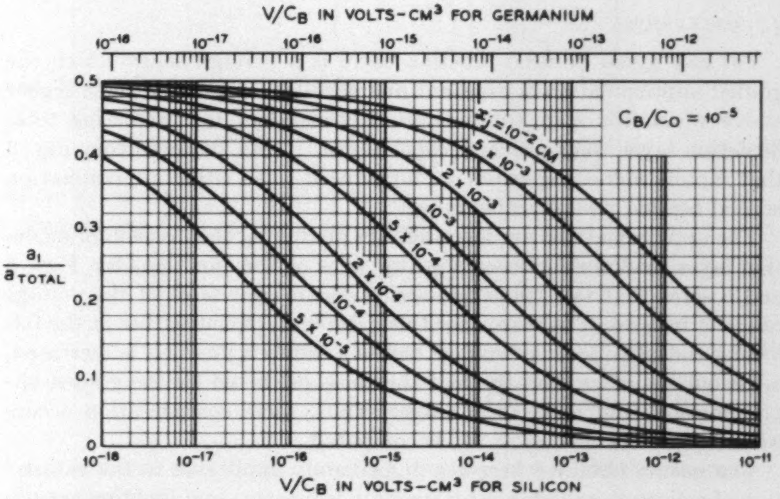

Fig. 6 — Chart for use in range $3 \times 10^{-6}$ to $3 \times 10^{-5}$, gaussian distribution.

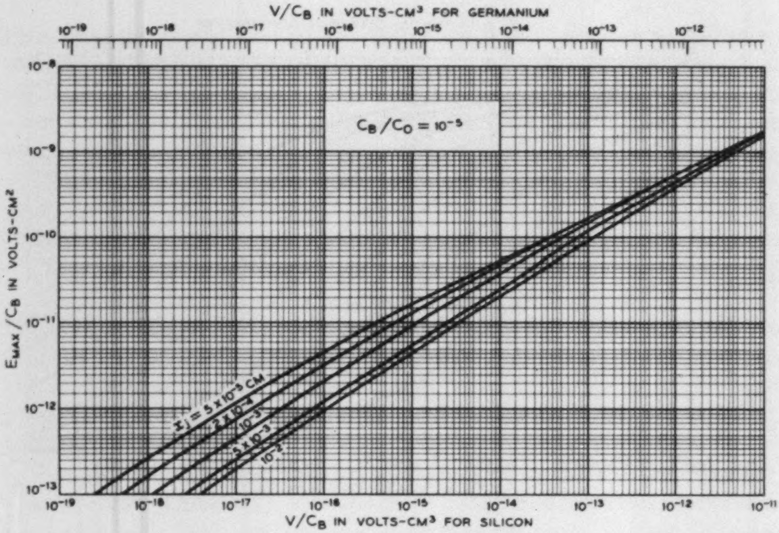

Fig. 7 — Peak electric field.

$C_0 = 10^{-6}$, with $C_B = 10^{14}$ and with errors of the same magnitude as mentioned above.

## V. CONCLUSIONS

For any given diffused junction there is a voltage below which the graded approximation is excellent and a voltage about a decade higher above which the step approximation is excellent for predicting total depletion layer thickness and capacitance. It can be seen from Fig. 2 that rapidly increasing errors are introduced when either approximation is used beyond its applicable range.

The approximations are less useful in predicting the division of depletion layer thickness between the two sides of the junction, for Figs. 5 and 6 show that this transition region occurs over most of the voltage range of interest. The reason for this difference, of course, lies in the following consideration: As voltage across a diffused junction is increased, $a_1$ becomes smaller than the half thickness predicted by the graded approximation, and $a_2$ becomes larger. Thus, some compensation occurs when total thickness, $a_1 + a_2$, is computed.

The results obtained here are immediately applicable to the calculation of collector capacitance for junction transistors and junction capacitor design and the like. In field effect device design, where single-side depletion layer thicknesses are important, the results are particularly helpful. Further, they should provide additional information for the detailed study of problems such as avalanche breakdown.

## VI. ACKNOWLEDGMENT

This work would not have been possible without the patient help provided by Mrs. Wanda Mammel of the mathematical research department, who prepared the error function subroutine for the IBM 704 and contributed a great deal towards the programming of our problem. Many thanks are due also to George Levenbach for many helpful discussions and for introducing us to the IBM 704. In addition, we wish to thank all the others whose interest and suggestions have contributed to this work.

## APPENDIX A

As shown in Section II, we must solve Poisson's equation, which in this case can be written

$$\nabla^2\psi = -\frac{qC_0}{\kappa\epsilon_0}\left(1 - \frac{C_B}{C_0} - \operatorname{erf}\frac{x}{\sqrt{4Dt}}\right). \tag{16}$$

All the symbols are defined as before. From (4) we have, for the electric field distribution on the left side of the junction,

$$E_1(x) = \frac{qC_0}{\kappa\epsilon_0}\int_{x_j-a_1}^{x}\left[\left(1 - \frac{C_B}{C_0}\right) - \operatorname{erf}\frac{x}{\sqrt{4Dt}}\right]dx, \tag{17}$$

and, from (5), for the right side,

$$E_2(x) = \frac{qC_0}{\kappa\epsilon_0}\int_{x_j+a_2}^{x}\left[\left(1 - \frac{C_B}{C_0}\right) - \operatorname{erf}\frac{x}{\sqrt{4Dt}}\right]dx, \tag{18}$$

where $x_j$, $a_1$ and $a_2$ are defined in Fig. 1.

Integrating (17) by parts, we obtain

$$\begin{aligned} E_1(x) = \frac{qC_0}{\kappa\epsilon_0}\Bigg\{&\left[\left(1 - \frac{C_B}{C_0}\right)x - x\operatorname{erf}\frac{x}{\sqrt{4Dt}} - \sqrt{\frac{4Dt}{\pi}}\,e^{-x^2/(4Dt)}\right] \\ &- \left[\left(1 - \frac{C_B}{C_0}\right)(x_j - a_1) - (x_j - a_1)\operatorname{erf}\frac{x_j - a_1}{\sqrt{4Dt}}\right. \\ &\left. - \sqrt{\frac{4Dt}{\pi}}\,e^{-(x_j-a_1)^2/(4Dt)}\right]\Bigg\}. \end{aligned} \tag{19}$$

Peak field will exist at the junction. Letting $x = x_j$ in this expression and noting that

$$1 - \frac{C_B}{C_0} = \operatorname{erf}\frac{x_j}{\sqrt{4Dt}}, \tag{20}$$

we can write for the peak field

$$E_1(x_j) = \frac{qC_0}{\kappa\epsilon_0}\left\{(x_j - a_1)\left(\operatorname{erf}\frac{x_j - a_1}{\sqrt{4Dt}} - \operatorname{erf}\frac{x_j}{\sqrt{4Dt}}\right) - \sqrt{\frac{4Dt}{\pi}}\,[e^{-x_j^2/(4Dt)} - e^{-(x_j-a_1)^2/(4Dt)}]\right\}. \tag{21}$$

For the range of arguments used in these calculations, the error function is very close to unity, e.g., 0.99999980. Because the IBM 704 can handle at most eight decimal digits, it is advantageous to work with the error function complement. Making this substitution, the final expression for peak field becomes

$$E_1(x_j) = \frac{qC_0}{\kappa\epsilon_0}\left\{(x_j - a_1)\left(\operatorname{erfc}\frac{x_j}{\sqrt{4Dt}} - \operatorname{erfc}\frac{x_j - a_1}{\sqrt{4Dt}}\right) - \sqrt{\frac{4Dt}{\pi}}\,[e^{-x_j^2/(4Dt)} - e^{-(x_j-a_1)^2/(4Dt)}]\right\}. \tag{22}$$

Similarly, (18) can be integrated to give an expression for the field distribution on the right-hand side of the junction and, specifically, the peak field:

$$E_2(x_j) = \frac{qC_0}{\kappa\epsilon_0}\left\{(x_j + a_2)\left(\operatorname{erfc}\frac{x_j}{\sqrt{4Dt}} - \operatorname{erfc}\frac{x_j + a_2}{\sqrt{4Dt}}\right) - \sqrt{\frac{4Dt}{\pi}}\,[e^{-x_j^2/(4Dt)} - e^{-(x_j+a_2)^2/(4Dt)}]\right\}. \tag{23}$$

The desired relation between $a_1$ and $a_2$ can be obtained by noting that field continuity requires

$$E_1(x_j) = E_2(x_j). \tag{24}$$

Thus, we obtain an equation in $a_2$,

$$F(a_2) = 0, \tag{25}$$

which will be solved by trial and error. From (22) and (23) it is evident that

$$F(a_2) = a_2 \operatorname{erfc}\frac{x_j}{\sqrt{4Dt}} - (x_j + a_2)\operatorname{erfc}\frac{x_j + a_2}{\sqrt{4Dt}} + \sqrt{\frac{4Dt}{\pi}}\,e^{-(x_j+a_2)^2/(4Dt)} - \left[(a_1 - x_j)\operatorname{erfc}\frac{x_j - a_1}{\sqrt{4Dt}} - a_1\operatorname{erfc}\frac{x_j}{\sqrt{4Dt}} + \sqrt{\frac{4Dt}{\pi}}\,e^{-(x_j-a_1)^2/(4Dt)}\right]. \tag{26}$$

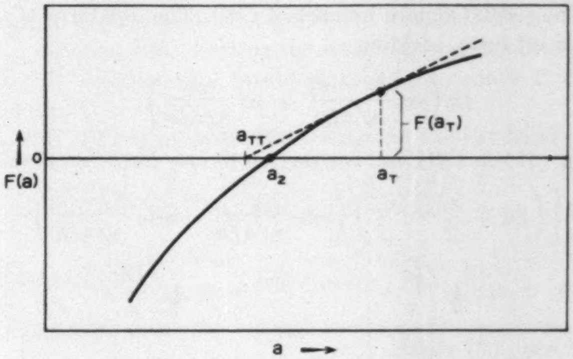

Fig. 8 — Method of obtaining approximation $a_{TT}$.

For rapid convergence on $a_2$ in the machine solution of (25), Newton's approximation was used.* For a trial value, $a_T$, which is a first approximation to $a_2$, we simply set $a_T = a_1$. By referring to Fig. 8 it can be seen that the next approximation, $a_{TT}$, can be obtained from

$$a_{TT} = a_T - \frac{F(a_T)}{\left.\dfrac{dF}{da}\right]_{a_T}}, \tag{27}$$

where

$$\left.\frac{dF}{da}\right]_{a_T} = \operatorname{erfc}\frac{x_j}{\sqrt{4Dt}} - \operatorname{erfc}\frac{x_j + a_T}{\sqrt{4Dt}}. \tag{28}$$

By substituting $a_{TT}$ back into (27) in place of $a_T$ (i.e., by repeated applications of Newton's approximation), rapid convergence results. A criterion of accuracy is the expression

$$\frac{E_1(x_j) - E_2(x_j)}{E_1(x_j)}, \tag{29}$$

where the approximate value of $a_2$ is used in evaluating $E_2(x_j)$.

To obtain the total voltage across the junction (applied + built-in), (8) and (9) of the text must be integrated. Equation (8) yields

$$\begin{aligned} V_1 = -\frac{qC_0}{\kappa\epsilon_0}\Bigg\{\Bigg[\left(1 - \frac{C_B}{C_0}\right)\frac{x^2}{2} - \frac{x^2}{2}\operatorname{erf}\frac{x}{\sqrt{4Dt}} \\ - \sqrt{\frac{Dt}{\pi}}\, x\, e^{-x^2/(4Dt)} - \kappa x\Bigg]_{x_j-a_1}^{x_j} - \sqrt{\frac{Dt}{\pi}}\int_{x_j-a_1}^{x_j} e^{-x^2/(4Dt)}\,dx\Bigg\}, \end{aligned} \tag{30}$$

---

* This was suggested by Miss M. C. Gray.

where $\kappa$ is the second square bracket of (19). The last term of (24) can be written in erf form, yielding

$$Dt\left(\operatorname{erf}\frac{x_j - a_1}{\sqrt{4Dt}} - \operatorname{erf}\frac{x_j}{\sqrt{4Dt}}\right). \tag{31}$$

Substituting (31) in (30) and converting to erfc form, we obtain

$$\begin{aligned} V_1 = -\frac{qC_0}{\kappa\epsilon_0}\Bigg[&\left(Dt + \frac{x_j^2}{2} - \frac{a_1^2}{2}\right)\left(\operatorname{erfc}\frac{x_j}{\sqrt{4Dt}} - \operatorname{erfc}\frac{x_j - a_1}{\sqrt{4Dt}}\right) \\ &+ (x_j + a_1)\sqrt{\frac{Dt}{\pi}}\,e^{-(x_j-a_1)^2/(4Dt)} - x_j\sqrt{\frac{Dt}{\pi}}\,e^{-x_j^2/(4Dt)}\Bigg]. \end{aligned} \tag{32}$$

In a similar way, (9) yields

$$\begin{aligned} V_2 = \frac{qC_0}{\kappa\epsilon_0}\Bigg[&\left(Dt + \frac{x_j^2}{2} - \frac{a_2}{2}\right)^2\left(\operatorname{erfc}\frac{x_j}{\sqrt{4Dt}} - \operatorname{erfc}\frac{x_j + a_2}{\sqrt{4Dt}}\right) \\ &+ (x_j - a_2)\sqrt{\frac{Dt}{\pi}}\,e^{-(x_j+a_2)^2/(4Dt)} - x_j\sqrt{\frac{Dt}{\pi}}\,e^{-x_j^2/(4Dt)}\Bigg]. \end{aligned} \tag{33}$$

## APPENDIX B

We must again obtain a solution for Poisson's equation, which in the case of the gaussian distribution can be written

$$\Delta^2\psi = \frac{qC_0}{\epsilon\epsilon_0}\left(e^{-x^2/(4Dt)} - \frac{C_B}{C_0}\right), \tag{34}$$

where the symbols are defined as before.

Integrations corresponding to those in Appendix A are carried out to obtain the electric field distribution on either side of the junction. These field functions are as follows:

$$\begin{aligned} E_1(x) = \frac{qC_0}{\kappa\epsilon_0}\Bigg[&\sqrt{\pi Dt}\left(\operatorname{erf}\frac{x}{\sqrt{4Dt}} - \operatorname{erf}\frac{x_j - a_1}{\sqrt{4Dt}}\right) \\ &- \frac{C_B}{C_0}(x - x_j + a_1)\Bigg] \end{aligned} \tag{35}$$

and

$$\begin{aligned} E_2(x) = \frac{qC_0}{\kappa\epsilon_0}\Bigg[&\sqrt{\pi Dt}\left(\operatorname{erf}\frac{x}{\sqrt{4Dt}} - \operatorname{erf}\frac{x_j + a_2}{\sqrt{4Dt}}\right) \\ &- \frac{C_B}{C_0}(x - x_j - a_2)\Bigg]. \end{aligned} \tag{36}$$

The expressions for peak field are found immediately by substituting $x_j$ for $x$ in (35) and (36). Setting the resulting expressions equal to each other gives an equation that establishes field continuity at the junction [(7) of the text].

Integrating (35) again gives the voltage on the left-hand side:

$$\begin{aligned} V_1 = -\frac{qC_0}{\kappa\epsilon_0}\Bigg\{&\sqrt{\pi Dt}\, x_j \left(\operatorname{erf}\frac{x_j}{\sqrt{4Dt}} - \operatorname{erf}\frac{x_j - a_1}{\sqrt{4Dt}}\right) \\ &+ 2Dt[e^{-x_j^2/(4Dt)} - e^{-(x_j-a_1)^2/(4Dt)}] - \tfrac{1}{2}a_1^2 e^{-x_j^2/(4Dt)}\Bigg\}. \end{aligned} \tag{37}$$

Integrating (36) gives the voltage on the right hand side:

$$\begin{aligned} V_2 = \frac{qC_0}{\kappa\epsilon_0}\Bigg\{&\sqrt{\pi Dt}\, x_j \left(\operatorname{erf}\frac{x_j}{\sqrt{4Dt}} - \operatorname{erf}\frac{x_j + a_2}{\sqrt{4Dt}}\right) \\ &+ 2Dt[e^{-x_j^2/(4Dt)} - e^{-(x_j+a_2)^2/(4Dt)}] - \tfrac{1}{2}a_2^2 e^{-x_j^2/(4Dt)}\Bigg\}. \end{aligned} \tag{38}$$

## REFERENCE

1. Barrer, R. M., *Diffusion In and Through Solids*, Macmillan, New York, 1941.

# A Transversal Equalizer for Television Circuits

By R. V. SPERRY and D. SURENIAN

(Manuscript received November 23, 1959)

*Transmission of television signals over long systems requires a fine degree of equalization. An adjustable type of equalizer to supplement fixed equalizers is discussed. It provides both gain and delay characteristics in the form of harmonically related cosine shapes which are independently adjustable. The design and operation of the equalizer is explained on the basis of paired echo theory. The 336A equalizer used in the L3 coaxial system is discussed as an example of this equalization technique.*

## I. INTRODUCTION

In an ideal transmission system the received signal is an exact replica of the transmitted signal. In all practical systems the received signal differs from the transmitted signal, due to imperfections of the transmission media. The lack of uniform gain and delay over the frequency band of the system constitutes a common form of imperfection. These gain and delay distortions are usually so severe as to require some form of equalization.* The amount of equalization required depends upon the nature of the signal being transmitted. Fortunately, most communication signals have qualities that make perfect equalization unnecessary.

For television transmission, the requirements are particularly stringent for both gain and delay equalization. Systems transmitting TV signals either as video circuits or carrier circuits usually use fixed gain and delay equalizers to correct the bulk of the system distortion and adjustable equalizers to correct the residual.

The residual gain and delay distortions will vary with system aging and manufacturing variations; thus, an unpredictable amount of equalization is necessary. These distortion characteristics are arbitrary functions of frequency, and will change from time to time. This requires that the adjustable equalizer be flexible enough to equalize arbitrary distortion characteristics, but still be easy to adjust.

---

* A theory of equalization of complex systems is described in Ref. 1.

It is well known that any reasonable function over a finite interval can be mathematically approximated by means of a finite number of terms of an infinite series. The degree of approximation determines the number of terms of the series. The Fourier cosine series is particularly useful because its orthogonal properties make the coefficients of each term independent. Furthermore, methods of producing cosine shapes of gain and delay are known. So a device that produces adjustable cosine shapes of gain and delay should be ideal for equalizing residual distortions. This paper describes an adjustable equalizer that produces harmonically related cosine shapes of both gain and delay. This device is called the *transversal equalizer.**

## II. THEORY OF THE TRANSVERSAL EQUALIZER

Wheeler's theory of paired echoes[4] is useful in describing the operation of the transversal equalizer. Let the impulse response of a network be the two impulses as shown by Fig. 1. The larger response impulse will be called the *main signal* and the smaller one the *echo*. As shown in Fig. 1, the response is a main signal and a lagging echo. The transfer function of the network can be obtained from the Laplace transform of the response, and is given by

$$\begin{aligned} F(j\omega) &= L[\delta(t_1) + K\delta(t_1 + \tau)] \qquad p = j\omega \\ &= e^{-jt_1\omega}(1 + Ke^{-j\tau\omega}) \\ &= e^{-(\alpha + j\beta)}. \end{aligned}$$

The loss of the network is given by

$$\alpha = -20 \log |F(j\omega)| \qquad \text{in decibels,}$$

and the phase is given by

$$\beta = -\tan^{-1} \frac{\mathrm{Im}\,[F(j\omega)]}{\mathrm{Re}\,[F(j\omega)]} \qquad \text{in radians,}$$

which, for $K \ll 1$, results in

$$\begin{aligned} \alpha &\doteq -8.686K \cos \tau\omega, \\ \beta &\doteq t_1\omega + K \sin \tau\omega \end{aligned} \tag{1}$$

---

* This name is a logical one in view of the use of "transversal filter" by Kallmann[2] and Linke.[3]

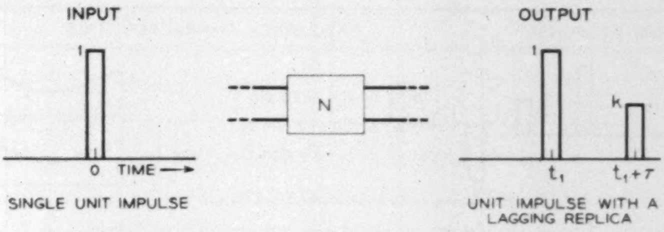

Fig. 1 — Single lagging echo response.

and the envelope delay

$$T \equiv \frac{d\beta}{d\omega} \doteq t_1 + \tau K \cos \tau\omega.$$

Thus, any network producing an echo pattern as shown by Fig. 1 has loss and delay responses in the frequency domain that are approximate cosines. Should the response be a main signal and a leading echo, the sign of $\tau$ would be negative and the resulting loss, phase and delay are given by

$$\begin{aligned} \alpha &\doteq -8.686K \cos \tau\omega, \\ \beta &\doteq t_1\omega - K \sin \tau\omega, \\ T &\doteq t_1 - K\tau \cos \tau\omega. \end{aligned} \tag{2}$$

Note that only the phase and delay are affected. Therefore, if the response of the network were to produce both leading and lagging echoes as well as a main signal the frequency characteristics would be the addition of (1) and (2). A number of echo patterns and their corresponding frequency responses are given by Fig. 2.

This shows that a pair of echoes with even symmetry about the main signal is associated with cosine loss deviations and constant delay, while echoes with odd symmetry are associated with cosine shapes of delay and approximately constant loss. Note also that the cosine function completes one-half cycle in a frequency band $B = 1/2\tau$.

Consider now a network with the multiple echo response as shown by Fig. 3. The frequency response is given (see the Appendix) by

$$\alpha \doteq -8.686\,[2K_1 \cos \tau\omega - 2K_2 \cos 2\tau\omega + 2\,K_3 \cos 3\tau\omega + \cdots],$$

$$\beta = t_1\omega$$

and

$$T = t_1 .$$

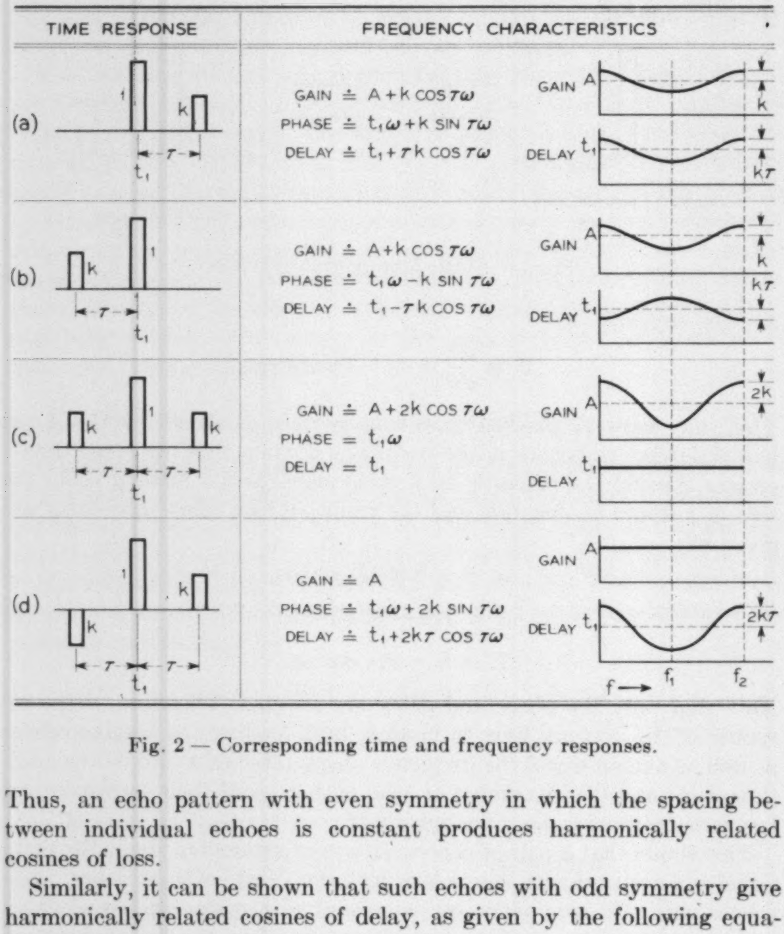
TIME RESPONSE
FREQUENCY CHARACTERISTICS
(a)
1
k
τ
t₁
GAIN ≐ A + k COS τω
PHASE ≐ t₁ω + k SIN τω
DELAY ≐ t₁ + τk COS τω
GAIN
A
k
DELAY
t₁
kτ
(b)
k
1
τ
t₁
GAIN ≐ A + k COS τω
PHASE ≐ t₁ω − k SIN τω
DELAY ≐ t₁ − τk COS τω
GAIN
A
k
kτ
DELAY
t₁
(c)
k
1
k
τ
τ
t₁
GAIN ≐ A + 2k COS τω
PHASE = t₁ω
DELAY = t₁
GAIN
A
2k
DELAY
t₁
(d)
1
k
k
τ
τ
t₁
GAIN ≐ A
PHASE ≐ t₁ω + 2k SIN τω
DELAY ≐ t₁ + 2kτ COS τω
GAIN
A
DELAY
t₁
2kτ
f →
f₁
f₂

Fig. 2 — Corresponding time and frequency responses.

Thus, an echo pattern with even symmetry in which the spacing between individual echoes is constant produces harmonically related cosines of loss.

Similarly, it can be shown that such echoes with odd symmetry give harmonically related cosines of delay, as given by the following equations:

$$\alpha \doteq K,$$

$$\beta \doteq t_1\omega + 2K_1 \sin \tau\omega + 2K_2 \sin \tau 2\omega + 2K_3 \sin \tau 3\omega + \cdots$$

and

$$T \doteq t_1 + 2K_1\tau \cos \tau\omega + 2K_2\,(2\tau) \cos \tau 2\omega + 2K_3\,(3\tau) \cos \tau 3\omega + \cdots.$$

Observe that the amplitudes of the harmonic terms of loss are propor-

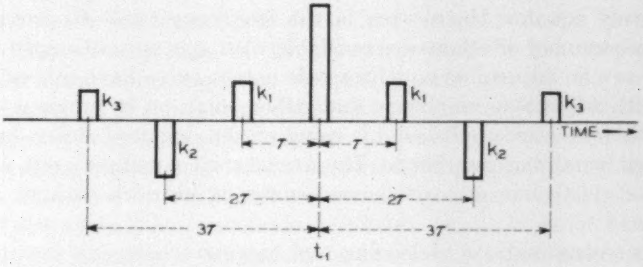

Fig. 3 — Multiple echo pairs with even symmetry.

tional to the amplitude of the echoes, while the amplitudes of the harmonic terms of delay are proportional to the product of the amplitudes of the echoes and the time interval between the corresponding echo and the main signal.

Networks capable of producing such multiple echo patterns have been used for Fourier series types of filters and equalizers by Wiener and Lee[5] and by others.[2,3,6,7]

A network capable of producing a multiple echo pattern can be constructed from a delay line with a number of equally spaced taps along the line, as shown by Fig. 4. The tap at the center of the line is for the so-called main signal. The other taps symmetrically spaced from the main signal tap are to provide leading and lagging echoes. In a restricted band $B$ the Nyquist sampling interval is $1/2B$; therefore the impulse response measured at $1/2B$ intervals is sufficient to describe the system completely. It follows that placing the taps $\tau$ seconds apart, where $\tau = 1/2B$ (which corresponds to a half cycle of the first cosine term), will

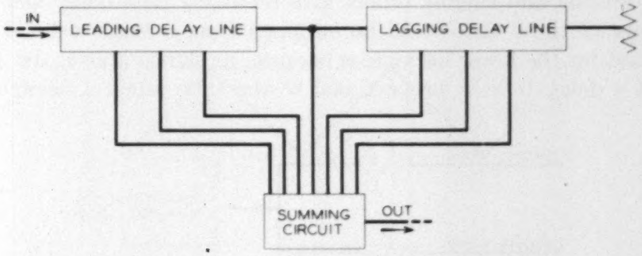

Fig. 4 — Block diagram of an echo-type equalizer.

completely equalize the system in the frequency band $B$, provided a sufficient number of echoes are available with appropriate amplitudes.

To have an adjustable equalizer it is necessary to have echoes available with adjustable amplitude and either polarity. In previous equalizers[3] this was accomplished by using either balanced delay lines or balanced summing amplifiers. The transversal equalizer uses an unbalanced delay line and a balanced summing network without active elements.[8]

Proper combinations of leading and lagging echoes can provide the desired harmonically related cosine shapes of loss only or delay only as shown by Fig. 2. The echoes can be combined electrically by introducing both a leading and a lagging echo to the same potentiometer, as shown in Fig. 5, if $r_1$ and $r_2$ are large compared to the potentiometer resistance. If it is necessary to have a single control adjust a pure loss term, both a leading and a lagging echo must be combined at the potentiometer corresponding to the desired term. These echoes should have the same amplitude and polarity to give a pure loss term. Now, if the same harmonic of delay is to be available and independently adjustable, the same two echoes must be introduced at the delay potentiometer. However, at this potentiometer one of the echoes must be of the opposite polarity to that introduced at the loss potentiometer. It follows then that, to adjust the equalizer in terms of loss and delay independently, both polarities of either the leading or lagging echoes will be simultaneously needed. This requires an additional delay line or a phase-inverting device. If independent adjustment is not necessary, a single echo can be used to adjust either loss or delay, and a pair of echoes can be used for the other. This would require that the single echo adjustment be made first and the paired echo adjustment be made later.

Some advantage accrues from using single echoes for the loss terms rather than the delay terms. Delay terms obtained from the combination of leading and lagging echoes give relatively pure delay and twice the range of those that could be obtained from a single echo. This is important for the lower harmonics because, as shown above, the amplitude of a delay term is proportional to the time interval between the

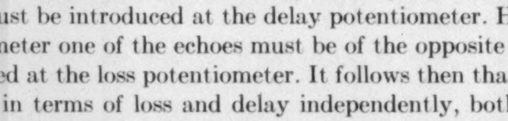

Fig. 5 — Method of combining leading or lagging echoes to produce pure loss or pure delay.

echo and the main signal, and hence proportional to the harmonic number. The disadvantage of this pattern is that the loss shapes also introduce a corresponding delay shape that may increase the over-all delay distortion. If the system delay distortion were of the minimum-phase type it would be advantageous to obtain the loss shapes from lagging echoes alone, since minimum-phase networks are lagging echo devices. However, if the system contains fixed-delay equalizers, the over-all distortion is not necessarily minimum phase. In order to secure optimum equalization, a switch may be provided to select either a leading or a lagging echo for each of the loss terms. This arrangement allows maximum utilization of the delay terms.

## III. ADJUSTING TECHNIQUE

The operation of the transversal equalizer can be described equally well in terms of paired-echo theory or in terms of its steady-state frequency response. These two descriptions, although equally correct, lead to two distinct methods of adjustment and evaluation of performance. The echo analysis leads to adjustment in terms of individual echoes and is described by Linke.[3] However, existing equipment suitable for adjusting the transversal equalizer utilizes a sweeping-frequency technique. This necessitates the evaluation of loss and delay characteristics separately, hence combinations of leading and lagging echoes are required.

The adjusting set is a modified version of that used for the "cosine equalizers."[1] Its operation is briefly described here in order to clarify the factors affecting the design of the transversal equalizer. For the loss adjustment, the adjusting set transmitter sends a swept-frequency signal through the system, as shown in Fig. 6. This signal is amplitude-modulated by the system loss distortion. The signal then travels through the equalizer and into the receiver of the adjusting set, where it is

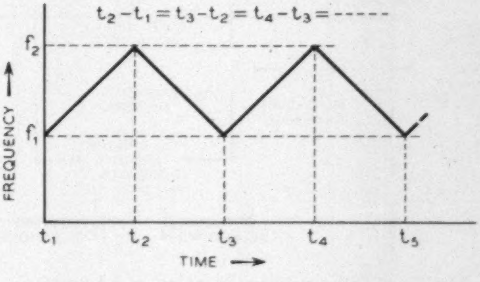

Fig. 6 — Sweep frequency of the cosine-adjusting set.

demodulated, and the resulting voltage is fed into a power meter. The waveform of this voltage represents the remaining distortion of the line plus the equalizer. Considering this distortion as a summation of its Fourier harmonics, it is apparent that removal of any harmonic by the equalizer reduces the total energy of the distortion voltage, and the power meter reading is lowered.

Delay adjustment is made similarly, except that two frequencies separated in frequency by a constant amount are swept across the band simultaneously. The difference in phase between these two frequencies at any instant is an indication of the delay characteristic of the line, and a discriminator converts this phase difference into a voltage which energizes the power meter.

Note that, in sweeping up and down the band, distortion versus frequency is converted into a periodic function of voltage versus time. It is also an even function about the time corresponding to either extreme of the sweep; therefore, to synthesize such a waveform, only Fourier cosine series terms are required. This restricts the equalizer to integral multiples of half cosines of frequency in the required band, which is consistent with the Nyquist rate. The up-and-down nature of the sweep converts these half-cycles into full cosine shapes in the repetition period.

## IV. PHYSICAL REALIZATION OF A TRANSVERSAL EQUALIZER

Several variations of the transversal equalizer have been constructed at Bell Telephone Laboratories. One variation is the 336A equalizer designed for use in the TV branch of the L3 coaxial system, as shown in Fig. 7. It provides 23 continuously adjustable terms of loss and 15 terms

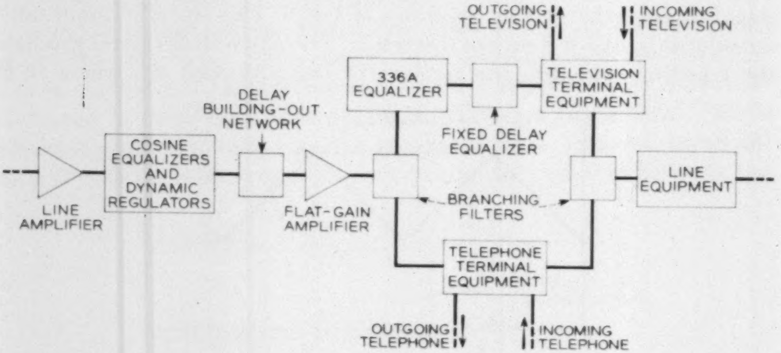
OUTGOING TELEVISION
INCOMING TELEVISION
336A EQUALIZER
TELEVISION TERMINAL EQUIPMENT
DELAY BUILDING-OUT NETWORK
FIXED DELAY EQUALIZER
COSINE EQUALIZERS AND DYNAMIC REGULATORS
LINE EQUIPMENT
LINE AMPLIFIER
FLAT-GAIN AMPLIFIER
BRANCHING FILTERS
TELEPHONE TERMINAL EQUIPMENT
OUTGOING TELEPHONE
INCOMING TELEPHONE

Fig. 7 — 336A equalizer location in L3 system.

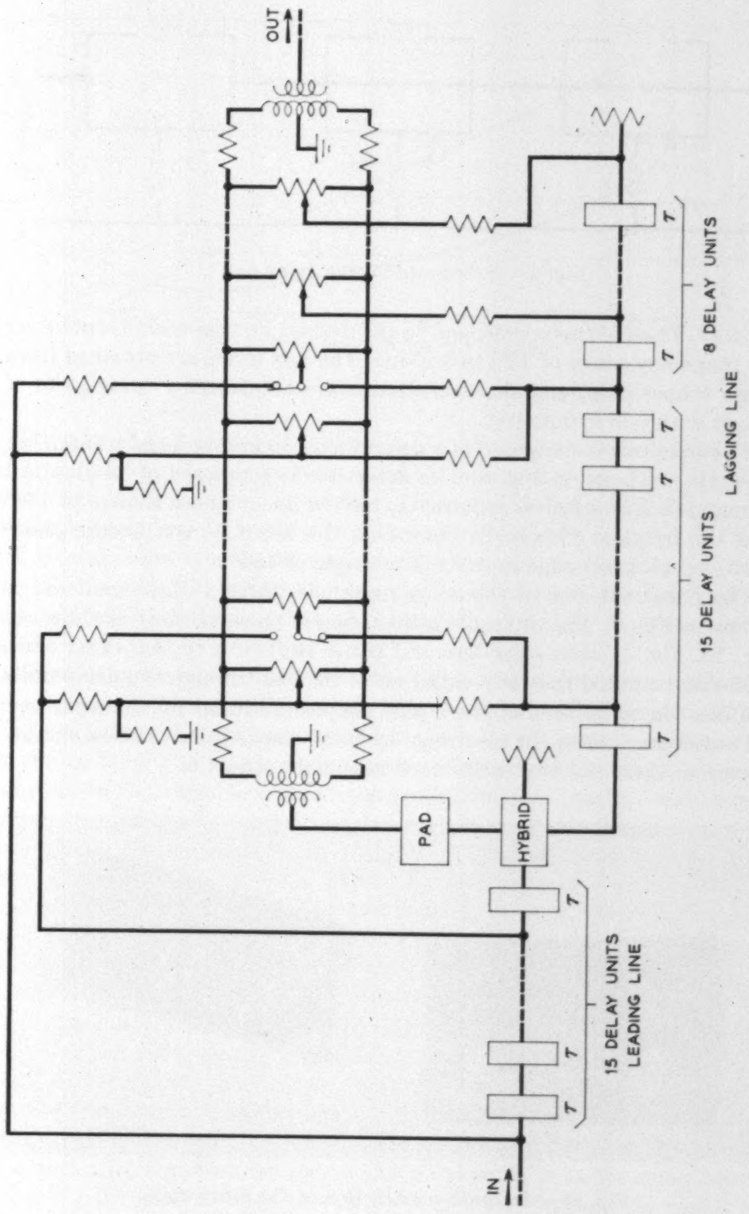
OUT
IN
PAD
HYBRID
T
T
T
T
T
T
T
T
15 DELAY UNITS
LEADING LINE
15 DELAY UNITS
LAGGING LINE
8 DELAY UNITS

Fig. 8 — Block schematic of transversal equalizer.

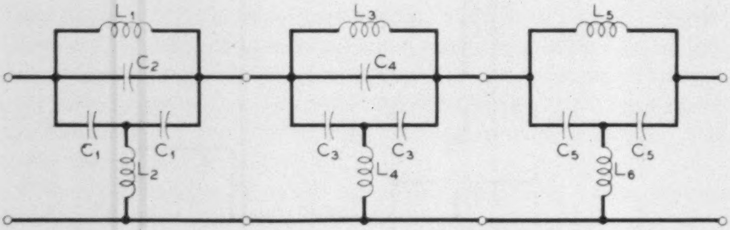

Fig. 9 — Schematic for the delay units.

of delay. These terms correspond to the desired Fourier cosine series over the frequency range of 3.75 to 8.25 mc. The loss terms are obtained from single echoes only, and the delay terms are obtained by using pairs of echoes with odd symmetry.

The equalizer is composed of a delay line, a hybrid coil and a summing network, as shown in Fig. 8. The delay line is composed of 38 identical sections, each of which is required to have a linear phase change of 180° over the frequency band. In order that the terms be cosines the phase shift at each band edge must be a multiple of 180°.

The present design of the delay units uses three all-pass sections, as shown in Fig. 9. The internal construction of these sections is shown in Fig. 10. Fig. 11 shows the loss and phase shift of a typical delay unit. It should be noted that any variation of the loss through the delay units modifies the amplitude of the cosine shapes produced by the equalizer. A Fourier analysis of the resulting "cosines" showed that the loss characteristic of these sections introduced negligible distortion.

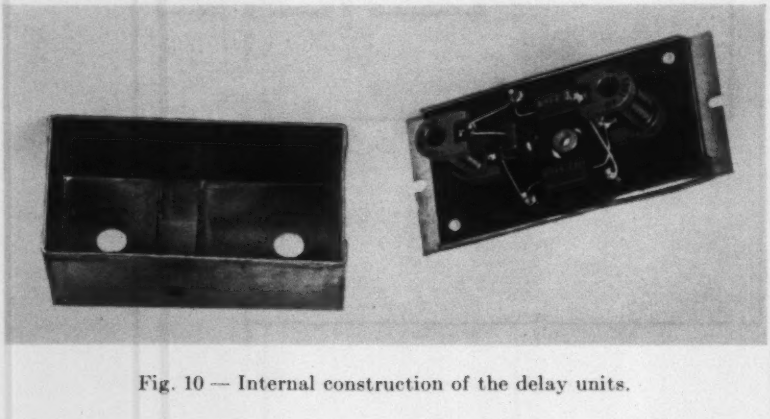

Fig. 10 — Internal construction of the delay units.

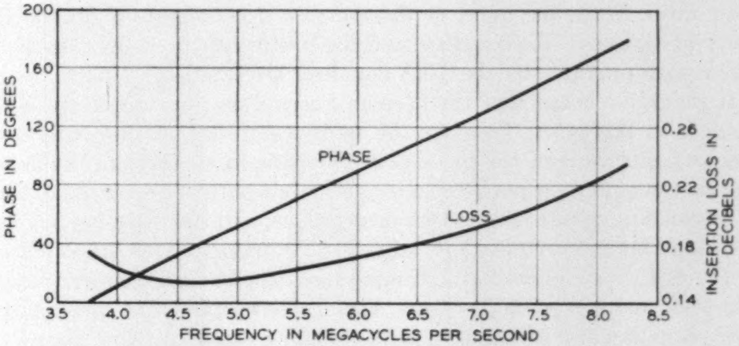

Fig. 11 — Loss and phase characteristics of a typical delay unit.

The hybrid coil serves several purposes. It splits the signal from the leading line into the component called the main signal and the component that energizes the lagging line. It inverts the phase of the signal to the lagging line thus providing the odd symmetry required for the delay terms. And it also provides isolation between the lagging line and the main signal; this is desirable because mismatches in the lagging line produce reflections that appear as lagging echoes if they are transmitted without attenuation to the output of the equalizer.

Combining the main signal with the tapped echoes requires a scheme that allows the magnitude and the sign of the echoes to be controlled. Using a balanced summing circuit obtained by means of transformers as shown by Fig. 8, the echo voltages are introduced through tapping resistors to the variable arms of the potentiometers that are connected across the balanced transformers. The position of the variable arm of the potentiometer controls the magnitude as well as the sign of the echo with respect to the main signal. The delay terms are obtained by combining corresponding leading and lagging echoes at the variable arm of a single potentiometer. In order that these echoes have equal intensities regardless of the setting of the potentiometer, a dissymmetrical pad is used at each delay potentiometer to compensate for the loss of the delay line and the hybrid coil.

To reduce the loading of the delay line at the tapped points, high values of resistors in the summing networks are required. This inherently restricts the echo magnitudes — and thus the ranges — unless the magnitude of the main signal is reduced. The pad between the hybrid and summing transformers serves this purpose; it is in the main transmission path and, if made arbitrarily large, will degrade the signal to noise ratio. Thus, the range of the loss and delay shapes is limited by the over-all loss of the equalizer and the loading of the line at the taps. The ranges provided by the 336A equalizer are listed in Table I.

It should be noted that the harmonic terms are not completely independent in the 336A. To make the terms completely orthogonal, very precise adjustment of the phase shift from the input through the delay line to the summing potentiometers for each path is required. A lack of precision here introduces term interactions, and consequently several iterations may be required to adjust the equalizer. This embodiment of the equalizer represents a compromise among cost factors, design schedules and ease of adjustment. Fig. 12 shows the front view of the 336A equalizer. It is equipped with a sliding mask to allow access to either the delay controls only or loss controls only. An additional knob is provided near each loss control to select leading or lagging echoes for loss equalization. The controls near the base are for simple all-pass networks that are switched in and out to supplement the first and second delay terms. Fig. 13 shows the equalizer with the sliding mask removed, and the controls for the selection of leading or lagging echoes can be seen.

## V. FIELD TRIAL

Models of the 336A equalizer were built at the Laboratories and tested in a 400-mile link of the L3 cable system. At the time of the tests, the

TABLE I

| $n$ | Loss Range, in db | Delay Range, in $\mu$sec |
|---|---|---|
| 1 | 1.1 | 0.03 |
| 2 | 1.1 | 0.06 |
| 3 | 1.1 | 0.09 |
| 4 | 1.1 | 0.12 |
| 5 | 1.1 | 0.14 |
| 6 | 0.8 | 0.09 |
| 7 | 0.7 | 0.10 |
| 8 | 0.7 | 0.11 |
| 9 | 0.7 | 0.12 |
| 10 | 0.7 | 0.13 |
| 11 | 0.5 | 0.11 |
| 12 | 0.5 | 0.12 |
| 13 | 0.5 | 0.12 |
| 14 | 0.5 | 0.13 |
| 15 | 0.5 | 0.13 |
| 16 | 0.4 | — |
| 17 | 0.4 | — |
| 18 | 0.4 | — |
| 19 | 0.4 | — |
| 20 | 0.4 | — |
| 21 | 0.4 | — |
| 22 | 0.4 | — |
| 23 | 0.4 | — |

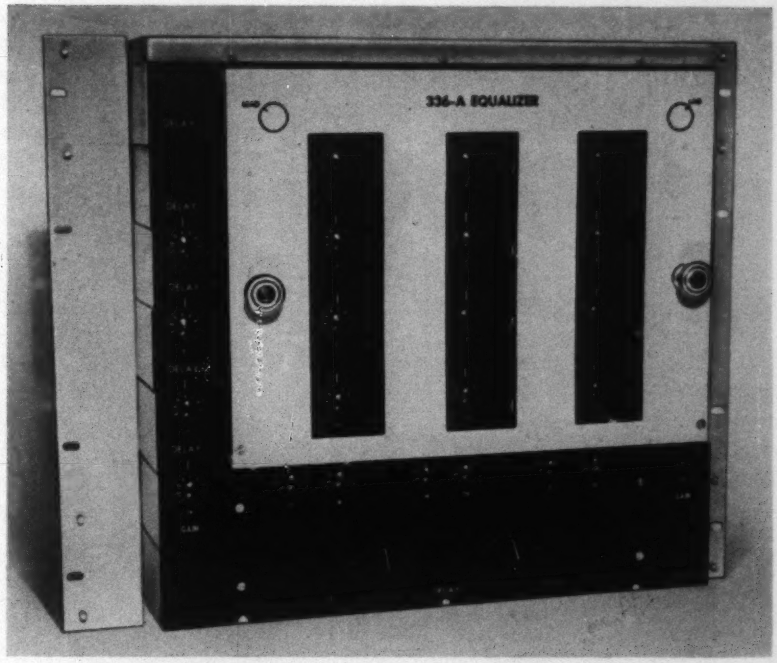

Fig. 12 — Front view of the 336A equalizer.

lines were equipped with only three of the full complement of six automatic regulators and therefore the system distortion cannot be considered that of a typical L3 system. The lines were otherwise in normal operating condition, including recent equalization of the entire L3 band with "cosine equalizers." Fig. 14(a) shows the gain and delay characteristics of a 400-mile line without a 336A equalizer. These transmission distortions do not represent any fixed type of deviations. In fact, the deviations from line to line are unpredictable, and therefore a very flexible equalizer such as the 336A is required. For this particular line the distortions of ±0.7 db and ±0.15 microseconds are excessive. However, Fig 14(b) shows the corresponding characteristic after adding a 336A equalizer and adjusting the 23 gain and the 15 delay terms. Note that the distortion in gain is now ±0.3 db over the entire band and the delay distortion is ±0.05 microsecond. The improvement in gain up to about 8 mc is quite marked (±0.1 db) and only at the extreme edge of the band does the gain deviate appreciably. It should be mentioned that the three regulators not in the system at the time these measure-

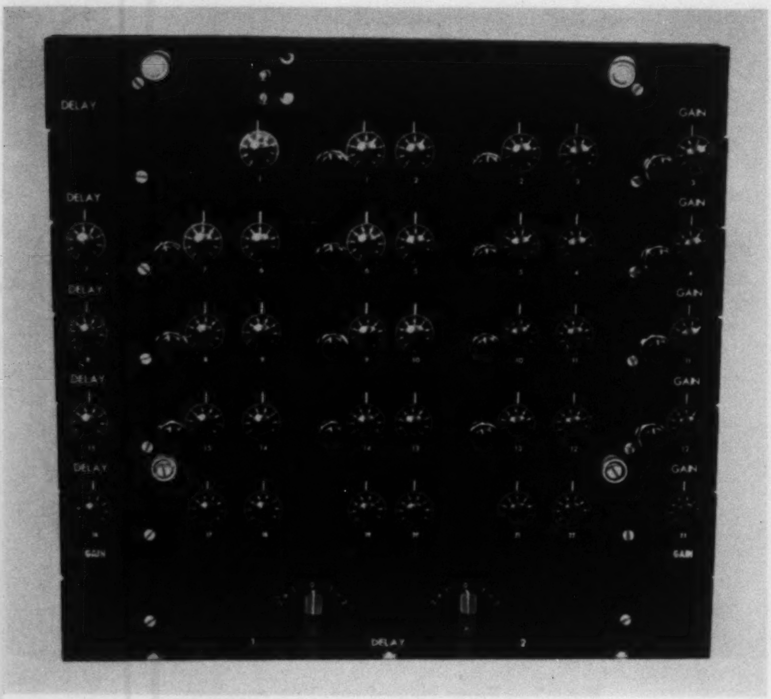
DELAY
GAIN
DELAY
GAIN
DELAY
GAIN
DELAY
GAIN
DELAY
GAIN
GAIN
GAIN
1
DELAY
2

Fig. 13 — 336A equalizer with sliding mask removed.

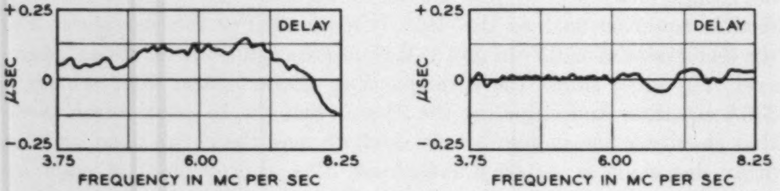

Fig. 14 — Delay and gain characteristics of a 400-mile circuit before and after equalization with a 336A equalizer.

ments were taken would be effective in reducing the gain distortion at the top edge of the band. Fig. 15 shows the gain and delay characteristic of a 800-mile loop after equalization with two 336A equalizers. Here the gain distortion is ±0.15 db and the delay distortion is ±0.07 microsecond over the desired band. The distortion over the 800-mile loop looks slightly better than the 400-mile line shown because the band edge characteristic happened to be less pronounced. There is no reason to expect that, in general, an 800-mile section with two equalizers would be better than a 400-mile section with one 336A.

## VI. CONCLUSIONS

The design of the transversal equalizer is a step forward in adjustable delay equalization. It is flexible because it uses the Fourier series approach; it can be made simple to adjust and is compatible with the existing adjusting techniques. It makes maximum use of only a few delay blocks to provide both loss and delay equalization. The flat loss of the equalizer is essentially independent of the number of terms; thus, it is particularly economical where a large number of terms is desired.

## VII. ACKNOWLEDGMENTS

The authors are indebted to many persons who have contributed to the realization of the 336A equalizer. Specifically, we wish to thank M. R. Aaron, F. J. Braga, R. Dempster, J. L. Garrison, R. S. Graham, E. S. Kuh, W. R. Lundry, O. L. Williams and G. F. Wyzga for their valuable suggestions and aid during the design and construction of the equalizer.

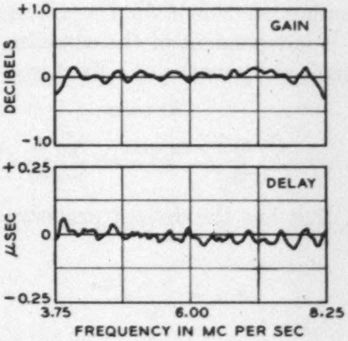

Fig. 15 — Delay and gain characteristics of a 800-mile circuit after equalization with two 336A equalizers.

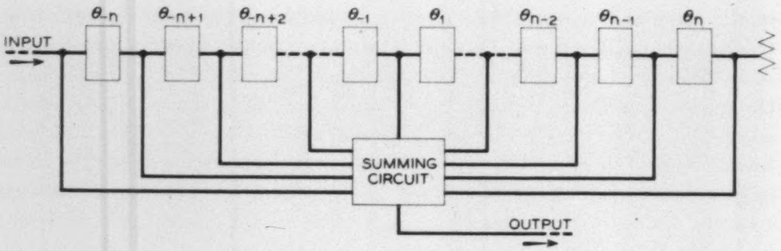

Fig. 16 — Block diagram of a transversal equalizer.

## APPENDIX

### *Multiple Echo Pattern Analysis*

Consider the performance of the circuit shown in Fig. 16, in which each box is a constant resistance network with

$$\frac{e_{\text{out}}}{e_{\text{in}}} = \epsilon^{\theta_k}$$

where

$$\theta_k = \alpha + j\beta \qquad k = 1, \cdots, n,$$

and both $\alpha$ and $\beta$ may be functions of frequency.

The insertion ratio of the circuit is given by

$$\frac{V_{\text{out}}}{V_{\text{in}}} = a_0 e^{n(\alpha+j\beta)} \left[ 1 + \sum_{k=1}^{n} \frac{a_{-k}}{a_0} e^{-k(\alpha+j\beta)} + \sum_{k=1}^{n} \frac{a_k}{a_0} e^{k(\alpha+j\beta)} \right],$$

where $a_i$ is the summing coefficient for the voltage from the $i$th box. The factor $a_0 e^{n(\alpha+j\beta)}$ is independent of the summing of the tapped voltages, and can be equalized by means of a fixed equalizer.

The factor

$$F = 1 + \sum_{k=1}^{n} \frac{a_{-k}}{a_0} e^{-k\alpha} e^{-j\beta k} + \sum_{k=1}^{n} \frac{a_k}{a_0} e^{+k\alpha} e^{j\beta k}$$

is the variable term that has the desired properties for equalization.

If we let

$$A_{-k} = \frac{a_{-k}}{a_0} e^{-k\alpha}$$

and

$$A_k = \frac{a_k}{a_0} e^{k\alpha},$$

the insertion gain variation of the circuit is given by

$$
\begin{aligned}
20 \log F = {} & \\
& 20 \log_{10}\left[1 + \sum_{k=1}^{n} (A_k + A_{-k}) \cos k\beta + j \sum_{k=1}^{n} (A_k - A_{-k}) \sin k\beta\right] \\
& \quad = G + j\Phi,
\end{aligned}
$$

where

$$
\begin{aligned}
G = {} & 10 \log_{10}\left\{1 + 2 \sum_{k=1}^{n} (A_k + A_{-k}) \cos k\beta \right. \\
& \left. + \left[\sum_{k=1}^{n} (A_k + A_{-k}) \cos k\beta\right]^2 + \left[\sum_{k=1}^{n} (A_k - A_{-k}) \sin k\beta\right]^2\right\} \qquad (3) \\
= {} & 10 \log_{10}\left[1 + 2 \sum_{k=1}^{n} (A_k + A_{-k}) \cos k\beta + D_1\right] \quad \text{decibels,}
\end{aligned}
$$

and

$$
\begin{aligned}
\Phi & = \arctan\left[\frac{\displaystyle\sum_{k=1}^{n} (A_k - A_{-k}) \sin k\beta}{1 + \displaystyle\sum_{k=1}^{n} (A_k + A_{-k}) \cos k\beta}\right] \\
& = \arctan\left[\sum_{k=1}^{n} A_k - A_{-k}) \sin k\beta - D_2\right],
\end{aligned} \qquad (4)
$$

where $D_1$ and $D_2$ are distortion terms containing cosines with magnitudes proportional to $A_m A_n$ and arguments of

$$
\left[(m \pm n)\beta + \frac{\delta\pi}{2}\right],
$$

where $\delta = 0, 1$.

Now, if we can assume $A_k < 1$, which means $a_k/a_o < 1$ (i.e., small percentage echo), then (3) becomes

$$
\begin{aligned}
G = 8.686 & \left\{\sum_{1}^{n} (A_k + A_{-k}) \cos k\beta - \left[\sum_{1}^{n} (A_k + A_{-k}) \cos k\beta\right]^2\right. \\
& \left. + \cdots + \frac{D_1}{2} - \frac{D_1^2}{4} + \cdots - D_1 \sum_{1}^{n} (A_k + A_{-k}) \cos k\beta + \cdots\right\}, \qquad (5)
\end{aligned}
$$

and (4) becomes

$$
\begin{aligned}
\Phi = {} & \sum_{1}^{n} (A_k - A_{-k}) \sin k\beta - D_2 \\
& - \frac{1}{3}\left[\sum_{1}^{n} (A_k - A_{-k}) \sin k\beta - D_2\right]^3.
\end{aligned} \qquad (6)
$$

If the $A_k$ are much smaller than unity, only the first term of either (5) or (6) need be considered. Thus the loss and phase are cosine and sine functions respectively only if these conditions are met.

Hence

$$G \doteq 8.686 \sum_{1}^{n} (A_k + A_{-k}) \cos k\beta,$$

$$\Phi \doteq \sum_{1}^{n} (A_k - A_{-k}) \sin k\beta.$$

REFERENCES

1. Ketchledge, R. W. and Finch, T. R., The L3 Coaxial System Equalization and Regulation, B.S.T.J., **32,** July 1953, p. 833.
2. Kallmann, H. E., Transversal Filters, Proc. I.R.E., **28,** July 1940, p. 302.
3. Linke, J. M., A Variable Time Equalizer for Video Frequency Waveform Correction, Proc. I.E.E., **99,** Part III, 1952, p. 427.
4. Wheeler, H. A., The Interpretation of Amplitude and Phase Distortion in Terms of Paired Echoes, Proc. I.R.E., **27,** June 1939, p. 359.
5. Wiener, N. and Lee, Y. W., U.S. Patents Nos. 2,024,900, December 1935; 2,124,599, July 1938; 2,128,257, August 1938.
6. Blumlein, A. D., Kallmann, H. E. and Percival, W. S., British Patent No. 517,516, June 1938.
7. Bellows, B. C. and Graham, R. S., Experimental Transversal Equalizer for TD-2 Radio Relay System, B.S.T.J., **36,** November 1957, p. 1429.
8. Graham, R. S. and Sperry, R. V., U.S. Patent No. 2,760,164, August 1956.

# Recent Monographs of Bell System Technical Papers Not Published in This Journal*

ANDERSON, E. W., see McCall, D. W.

ASCHNER, J. F., BITTMANN, C. A., HARE, W. F. J. and KLEIMACK, J. J.

**Silicon High-Frequency Switching Transistor Produced by Masking,** Monograph 3312.

BERRY, R. W. and SLOAN, D. J.

**Tantalum Printed Capacitors,** Monograph 3313.

BITTMANN, C. A., see Aschner, J. F.

COLLIER, R. J., see Feinstein, J.

CORENZWIT, E., see Nesbitt, E. A.

DARLINGTON, S.

**Nonstationary Smoothing and Prediction Using Network Theory Concepts,** Monograph 3350.

DAVID, E. E., JR., GUTTMAN, N. and VAN BERGEIJK, W. A.

**Binaural Interaction of High-Frequency Complex Stimuli,** Monograph 3392.

DILLON, J. F., GESCHWIND, S., JACCARINO, V. and MACHALETT, A.

**Cryostat for Ferrimagnetic Resonance Experiments,** Monograph 3394.

DOUGLASS, D. C., see McCall, D. W.

DUNN, H. K., see Sivian, L. J.

---

* Copies of these monographs may be obtained on request to the Publication Department, Bell Telephone Laboratories, Inc., 463 West Street, New York 14, N. Y. The numbers of the monograhs should be given in all requests.

FEINSTEIN, J. and COLLIER, R. J.

**A Class of Waveguide-Coupled Slow-Wave Structures,** Monograph 3319.

FINZI, L. A. and SUOZZI, J. J.

**On Feedback in Magnetic Amplifiers,** Monograph 3320.

FLANAGAN, J. L.

**Estimates of Intraglottal Pressure During Phonation,** Monograph 3359.

GESCHWIND, S., see Dillon, J. F.

GUTTMAN, N., see David, E. E., Jr.

HANSEN, R. H., see Hawkins, W. L.

HARE, W. F. J., see Aschner, J. F.

HASZKO, S. E., see Van Uitert, L. G.

HAWKINS, W. L., HANSEN, R. H., MATREYEK, W., WINSLOW, F. H., LANZA, V. L. and LOEFFLER, B. B.

**Carbon Black and New Thermal Antioxidants for Polyethylene,** Monograph 3321.

HENSEL, J. C. and PETER, M.

**Stark Effect for Cyclotron Resonance in Degenerate Bands,** Monograph 3384.

JACCARINO, V., see Dillon, J. F.

KLEIMACK, J. J., see Ascher, J. F.

KLEIN, J. M., see Thomas, D. E.

KRAMER, H. P.

**A Generalized Sampling Theorem,** Monograph 3323.

LANZA, V. L., see Hawkins,, W. L.

LAW, J. T.

**Adsorption of Hydrogen on Silicon,** Monograph 3391.

LOEFFLER, B. B., see Hawkins, W. L.

Looney, D. H.

**Recent Advances in Magnetic Devices for Computers,** Monograph 3367.

Lovell, L. C. and Wernick, J. H.

**Dislocation Etch Pits in Bismuth and High-Purity Copper,** Monograph 3347.

Machalett, A., see Dillon, J. F.

Mason, W. P.

**Adhesion Between Metals and Its Effect on Fixed and Sliding Contacts,** Monograph 3368.

Mason, W. P.

**Energy Conversion in the Solid State IV,** Monograph 3324.

Mason, W. P.

**Internal Friction, Plastic Strain and Fatigue in Metals and Semiconductors,** Monograph 3326.

Matreyek, W., see Hawkins, W. L.

McCall, D. W., Douglass, D. C. and Anderson, E. W.

**Diffusion in Ethylene Polymers,** Monograph 3325.

Miller, R. C. and Savage, A.

**Asymmetric Hysteresis Loops and the Pyroelectric Effect in Barium Titanate,** Monograph 3370.

Miller, R. L.

**Nature of the Vocal Cord Wave,** Monograph 3393.

Nesbitt, E. A., Wernick, J. H. and Corenzwit, E.

**Magnetic Moments of Alloys and Compounds of Rare Earth-Iron Alloys,** Monograph 3327.

Peter, M., see Hensel, J. C.

Ross, I. M.

**Switching Transistors,** Monograph 3329.

Sanders, T. M., see Weinreich, G.

SAVAGE, A., see Miller, R. C.

SIVIAN, L. J., DUNN, H. K. and WHITE, S. D.

**Absolute Amplitudes and Spectra of Certain Musical Instruments and Orchestras,** Monograph 3381.

SLICHTER, W. P.

**Crystal Structures in Polyamides Made from ω-amino Acids,** Monograph 3354.

SLICHTER, W. P.

**Molecular Characteristics of Rubber-Like Materials,** Monograph 3330.

SLOAN, D. J., see Berry, R. W.

SPITZER, W. G. and WHELAN, J. M.

**Infrared Absorption and Electron Effective Mass in n-Type Gallium Arsenide,** Monograph 3382.

SUOZZI, J. J., see Finzi, L. A.

SWANEKAMP, F. W., see Van Uitert, L. G.

THOMAS, D. E. and KLEIN, J. M.

**Automatic Transistor Alpha Measuring Set,** Monograph 3331.

VAN BERGEIJK, W. A., see David, E. E., Jr.

VAN UITERT, L. G., SWANEKAMP, F. W. and HASZKO, S. E.

**Line Widths in Polycrystalline Yttrium Iron Garnet,** Monograph 3332.

WALTZ, M. C.

**Microwave Impedance Measurement in Junction Region of a Semiconductor,** Monograph 3380.

WEINREICH, G., SANDERS, T. M., JR. and WHITE, H. G.

**Acoustoelectric Effect in n-Type Germanium,** Monograph 3383.

WERNICK, J. H., see Lovell, L. C.

WERNICK, J. H., see Nesbitt, E. A.

WHELAN, J. M., see Spitzer, W. G.

WHITE, D. L.

**β-Quartz as High-Temperature Piezoelectric Material,** Monograph 3360.

WHITE, H. G., see Weinreich, G.

WHITE, S. S., see Sivian, L. J.

WINSLOW, F. H., see Hawkins, W. L.

# Contributors to this Issue

A. Busala, Dr. in E.E., 1945, University of Rome (Italy); M.S., 1951, Catholic University of America; research associate, Catholic University of America, 1948–53; Bell Telephone Laboratories, 1953 —. He has been engaged in fundamental development of telephone circuits, with special emphasis on the design of transistorized telephone sets for electronic switching systems. Member Sigma Xi.

Harold E. Curtis, B.S. and M.S., 1929, Massachusetts Institute of Technology; American Telephone & Telegraph Company, 1929–34; Bell Telephone Laboratories, 1934 —. He was concerned with early transmission studies of the coaxial cable and waveguide, and has specialized in transmission engineering studies of multichannel carrier telephony and microwave radio relay systems.

U. F. Gianola, B.Sc., 1948, and Ph.D., 1951, University of Birmingham, (England); Royal Aircraft Establishment, 1951; postdoctoral fellow, University of British Columbia, 1951–53; Bell Telephone Laboratories, 1953 —. As a member of the transmission research department he took part in experimental and theoretical studies of transmission line structures, analyses of a new magnetostrictive transducer, solar battery application to communications channels and fundamental studies of the effects of ion bombardment. At present he is engaged in studies of solid state memory and logic devices. Member American Physical Society, Research Society of America; senior member I.R.E.

E. N. Gilbert, B.S., 1943, Queens College; Ph.D., 1948, Massachusetts Institute of Technology; M.I.T. Radiation Laboratory, 1944–46; Bell Telephone Laboratories, 1948 —. Mr. Gilbert has been engaged in studies of information theory and switching theory. Recipient M.I.T. Applied Mathematics Fellowship, 1946–48. Member American Mathematical Society.

H. J. Landau, A.B., 1953, and Ph.D., 1957, Harvard University; teaching fellow, Harvard, 1956–57; Bell Telephone Laboratories, 1957 —.

He was engaged in mathematical research in function theory and linear spaces. Mr. Landau is now on leave of absence while studying at the Institute for Advanced Study. Member American Mathematical Society, Phi Beta Kappa.

H. **LAWRENCE**, B.S., 1949, Randolph-Macon College; M.S., 1950, Columbia University Teachers College; Tung-Sol Electric, 1954–56; Bell Telephone Laboratories, 1956 —. He has been engaged in development of semiconductor devices and components.

L. A. **MACCOLL**, A.B., 1919, University of Colorado; A.M., 1925, and Ph.D., 1934, Columbia University; Western Electric Co., 1919–25; Bell Telephone Laboratories, 1925 —. He has been engaged in general mathematical research and consultation. He is the author of the book, *Fundamental Theory of Servomechanisms*, and has been a visiting lecturer at Princeton University and adjunct professor of mathematics at New York University. Fellow New York Academy of Sciences; member American Mathematical Society, American Physical Society, Edinburgh Mathematical Society, Mathematical Association of America, Phi Beta Kappa, Sigma Xi.

H. O. **POLLAK**, B.A., 1947, Yale University; M.A., 1948, and Ph.D., 1951, Harvard University; Bell Telephone Laboratories, 1951 —. He has been engaged in mathematical analysis of gunnery and missile systems and mathematical research in communications. Member American Mathematical Society, Mathematical Association of America, School Mathematics Study Group.

R. V. **SPERRY**, B.S.E.E., 1949, and M.S.E.E., 1951, West Virginia University; Bell Telephone Laboratories, 1952 —. He has been concerned with filters, equalizers and other networks for television transmission over coaxial cable, microwave and other systems. At present he supervises a group working on development of repeaters for waveguide transmission systems. Member A.I.E.E., Eta Kappa Nu, Pi Tau Sigma, Tau Beta Pi.

D. **SURENIAN**, B.S.E.E., 1950, Syracuse University; M.S.E.E., 1952, Clarkson College of Technology; Bell Telephone Laboratories, 1952–57. His work at Bell Laboratories was concerned with the development of networks and filters for communications systems.

**Raymond M. Warner, Jr.**, B.S., 1947, Carnegie Institute of Technology; M.S., 1950, and Ph.D., 1952, Case Institute of Technology; Bell Telephone Laboratories, 1952–59. Mr. Warner's work at Bell Laboratories was in semiconductor device development. Member American Association for the Advancement of Science, Phi Kappa Phi, Sigma Xi, Tau Beta Pi.